Average-Consensus Algorithms in a Deterministic Framework

Kevin Topley, Vikram Krishnamurthy

Department of Electrical and Computer Engineering

The University of British Columbia, Vancouver, Canada

Email: {kevint, vikramk}@ece.ubc.ca

Index Terms

distributed algorithm, linear update, recurring connectivity, least-squares problem, average-consensus

Abstract

We consider the average-consensus problem in a multi-node network of finite size. Communication between nodes is modeled by a sequence of directed signals with arbitrary communication delays. Four distributed algorithms that achieve average-consensus are proposed. Necessary and sufficient communication conditions are given for each algorithm to achieve average-consensus. Resource costs for each algorithm are derived based on the number of scalar values that are required for communication and storage at each node. Numerical examples are provided to illustrate the empirical convergence rate of the four algorithms in comparison with a well-known "gossip" algorithm as well as a randomized information spreading algorithm when assuming a fully connected random graph with instantaneous communication.

Glossary

- BM, Bench-Mark Algorithm, proposed here in $(10) - (14)$.
- DA, Distributed-Averaging Algorithm, proposed here in $(15) - (19)$.
- OH, One-Hop Algorithm, proposed here in $(20) - (24)$.
- DDA, Discretized Distributed-Averaging, proposed here in $(27) - (32)$.
- Gossip, Gossip Algorithm, proposed in [3], also defined in $(156) - (158)$.
- RIS, Randomized Information Spreading, proposed in [19].
- ARIS, Adapted Randomized Information Spreading, defined in $(160) - (166)$.
- S$\mathcal{V}$SC, a "singly $\mathcal{V}$-strongly connected" communication sequence, defined in (34).
- S$\mathcal{V}$CC, a "singly $\mathcal{V}$-completely connected" communication sequence, defined in (36).
- I$\mathcal{V}$SC, an "infinitely $\mathcal{V}$-strongly connected" communication sequence, defined in (35).

- I$\mathcal{V}$CC, an "infinitely $\mathcal{V}$-completely connected" communication sequence, defined in (37).

I. INTRODUCTION

Average-consensus formation implies that a group of distinct nodes come to agree on the average of their initial values, see [3], [7], [10], [19], [21], [29], [6] for related work. Past results indicate that obtaining average-consensus on a set of of n initial vectors each with dimension d requires at least one of the following assumptions:

- storage and communication of a set with cardinality upper bounded by $O(nd)$ at each node, e.g. the "flooding" method described for example in [4], [16] (where O denotes the Landau big Oh)

- construction of directed acyclical graphs in the network topology [6], [1], [16]

- knowledge at the transmitting node of the receiving node's identity [17], [6], [1], [16]

- instantaneous communication as well as knowledge at the transmitting node of its out-degree [12], [9]

- strictly bi-directional and instantaneous communication [5], [8], [7], [10], [11], [29]

- instantaneous communication and symmetric probabilities of node communication [10], [12]

- an approximate average-consensus based on randomized number generation [19], [6], [1]

- pre-determined bounds on the communication delays as well as the use of averaging-weights that are globally balanced and pre-determined off-line [28], [15], [14], [13], [25], [21].

This paper proposes four algorithms to solve the average-consensus problem under weaker communication conditions than those listed above. We denote the four algorithms *Bench-Mark (BM), Distributed-Averaging (DA), One-Hop (OH)*, and *Discretized Distributed-Averaging (DDA)* .

(i) The BM algorithm is based on the "flooding" method described in [4], [16]. We show in Theorems 4.2, 4.3 that the BM algorithm achieves average-consensus given the weakest communication condition necessary for average-consensus under any distributed algorithm (this is the S$\mathcal{V}$SC condition, defined in Sec.IV).

(ii) As the main result, we show in Theorem 4.5 that the DA algorithm (a reduction of the BM algorithm) can achieve average-consensus under a *recurring* S$\mathcal{V}$SC condition (this is the I$\mathcal{V}$SC condition, defined in Sec.IV). By "recurring" we mean that for an infinite set of disjoint time intervals, the S$\mathcal{V}$SC condition occurs on each interval. Previous results based on iterative averaging (e.g. [3], [10], [26], [21], [25]) require special cases of the I$\mathcal{V}$SC condition.

(iii) The OH and DDA algorithms can be viewed respectively as simplified versions of the BM and DA algorithms. We will show that analogous results hold under these algorithms with respect to the S$\mathcal{V}$CC and I$\mathcal{V}$CC conditions defined in Sec.IV.

In contrast to earlier work, the main results of this paper show that under general uni-directional connectivity conditions on the communication sequence, each proposed algorithm achieves average-consensus in the presence of

- arbitrary communication delays,

- arbitrary link failures.

Another distinct contrast each of the proposed algorithms has with previously considered consensus algorithms is,

- each node will know exactly when the true average-consensus estimate has been locally obtained, regardless of the communication pattern between nodes.

Of course our general results come at a price. The main drawback of three of our proposed algorithms (DA, DDA, OH) is that they require local storage and transmission upper bounded by $O(n + d)$, where we recall that n is the network size and d is the dimension of the initial consensus vectors. Most average-consensus algorithms in the literature require only $O(d)$ costs, however they also require assumptions such as instantaneous communication, pre-determined averaging weights, or control at the transmitting node of the receiver identity. Our algorithms (DA, DDA, OH) require none of these assumptions and are particularly advantageous for average-consensus involving $d \geq n$ distinct scalars. In this case the transmission and storage cost of each algorithm is $O(d)$, hence the assumption of relatively weak communication conditions can be leveraged against past algorithms. There are several examples in the literature where $d \geq n$. For instance, if each node observes a $\sqrt{n}$ dimensional process or parameter in noise, then a distributed Kalman filter or maximum-likelihood estimate requires an average-consensus on $d = n + \sqrt{n}$ scalars (see [20], [4]).

The first algorithm we consider (BM) is an obvious solution and is presented here only because (i) it serves as a bench-mark for the other algorithms (ii) the communication conditions necessary for its convergence will be used in our main results, and (iii) its formal description has many of the same properties as the other three proposed algorithms. The BM algorithm requires local storage and transmission of $O(nd)$ and has optimal convergence rate (see Theorem 4.2-4.3, in Sec.IV).

A. Relation to Past Work

Since the literature in consensus formation is substantial, we now give an overview of existing results and compare them with the results of this paper. For sufficiently small $\epsilon > 0$, if $x_k = [x_k^1, x_k^2, \ldots, x_k^n] \in \mathbb{R}^n$ denotes the network state at time k, then [21] proves that each element in the sequence $\{x_k = (I - \epsilon\mathcal{L}_k)x_{k-1} , \ k = 1, 2, \ldots\}$ converges asymptotically to $\sum_{i=1}^{n} x_0^i/n$ if each graph Laplacian $\mathcal{L}_k$ in the sequence $\{\mathcal{L}_k \in \mathbb{R}^{n \times n} : \ k = 1, 2, \ldots\}$ is balanced and induces a strongly connected graph. The work [26] generalizes this result by allowing ϵ to decrease at a sufficiently slow rate, and assuming only that there exists some integer β such that the union of all graph Laplacians over the interval $[k\beta, (k + 1)\beta - 1]$ induces a strongly connected graph for each $k \geq 1$. However, each graph Laplacian in [26] is still assumed to be balanced, and as is typical of many interesting papers on average-consensus, neither [21] nor [26] explain any method by which the nodes can *distributively* ensure each $\mathcal{L}_k$ is balanced *while* the sequence $\{\mathcal{L}_k : \ k = 1, 2, \ldots\}$ is occurring. Hence the results of these works assume that all averaging weights are globally pre-determined off-line, or in other words, every node in the network is assumed to know what averaging weights they should locally be using at each iteration to guarantee the resulting $\mathcal{L}_k$ is globally balanced.

In contrast, [5] proposes a "Metropolis" algorithm that requires a "two-round" signaling process wherein nodes can *distributively* compute local averaging weights. It is shown in [5] that each element in the sequence $\{x_k = (I - \epsilon\mathcal{L}_k)x_{k-1}\ ,\ k = 1, 2, \ldots\}$ converges asymptotically to $\sum_{i=1}^{n} x_0^i/n$ under mild connectivity conditions on the sequence $\{\mathcal{L}_k\ :\ k = 1, 2, \ldots\}$. The work [9] also proposes a distributed algorithm that does not require pre-determined averaging weights, however [9] assumes each transmitting node knows the number of nodes that will receive its message, even before the message is transmitted. Similarly, the algorithm in [5] assumes bi-directional communication, and furthermore each of the stated results in [21], [26], [5], [9] assume that the communication is instantaneous. In contrast, only [21], [9] and [5] assume the communication is noiseless.

The results in this paper do not assume instantaneous or bi-directional communication, nor do they assume that the transmitting node knows what node will receive their message or when. However, our results do require noiseless communication. The issue of noisy communication can be treated as a "meta-problem" that may be super-imposed upon the framework considered here. Similar to our current approach, there has been much research that assumes noiseless communication (for instance [12], [17], [19]). Conversely, there is a growing body of work that considers average-consensus formation in specifically noisy communication settings [11], [26], [14].

Works such as [3], [18], [27], [15], [24], [2], [23], [21], [25], [15] require node communication properties that are special cases of the I$\mathcal{V}$SC condition defined in Sec.IV-A. However, besides the flooding BM method and the DA algorithm proven to converge here, the only other known algorithm that can even be conjectured to almost surely obtain average-consensus under all I$\mathcal{V}$SC sequences is a specific adaptation of the randomized information spreading (RIS) algorithm proposed in [19]. Our adaptation of this algorithm is referred to as ARIS and is detailed in Sec.VII-B of Appendix VII. We note, however, that the lower bound on convergence rate derived in [19] assumes instantaneous bi-directional communication, and furthermore, any version of the RIS algorithm assumes that the initial consensus variables are all positive valued. The Gossip algorithm proposed in [3] as well as the ARIS algorithm are used as points of reference for the four algorithms proposed in this paper. In Sec.V-A we compare the resource costs of all six algorithms, and in Sec.V-B the performance of each algorithm is illustrated by simulation under various randomized communication sequences when assuming a full network graph.

We note that [18], [27], [24], [2], [23] do not guarantee the final consensus value will equal the initial average; [3] proposes an update rule that assumes instantaneous bi-directional communication; and [25], [15], [21] assume pre-determined bounds on the communication delays as well as averaging-weights that are globally balanced at each iteration (see Theorem 5, [25]). In contrast, Theorem 4.5 in Sec.IV states that the DA algorithm will guarantee average-consensus under *all* I$\mathcal{V}$SC sequences, regardless of the communication delays and without requiring any pre-determined balancing of the averaging weights. We note that a distinct feature that all four of the proposed algorithms have is that they each require the local initial consensus vector to be stored in the respective database of each node until average-consensus is locally obtained at that node. Without this property, each algorithm could still ensure a "consensus formation" under the exact same communication conditions assumed in our main results, however the final consensus value would not necessarily equal the initial average, which is desirable in most applications [3], [4], [6], [1], [16], [20]. In further contrast to past results, the proofs of convergence used in this

work do not rely on matrix theory [5], [21], Lyapunov techniques [18], [28], or stochastic stability [10], [25], [13]; instead, for our main results we obtain a variety of lower bounds on the "error" reduction and show that under (deterministic) recurring connectivity conditions an average-consensus will asymptotically obtain in the L^2 norm. As a final note, we clarify that the communication is assumed to be causal, hence a signal cannot be received at node i before it has left node j. Given this assumption, our framework considers *every* possible sequence of signals, hence any realization of a (causal) stochastic communication model is a special case of our deterministic framework.

B. Outline

Sec.II formulates the problem statement as well as our assumptions regarding the node communication and algorithm framework. Sec.III defines the four proposed algorithms. Sec.IV states the communication conditions that are necessary and sufficient for each algorithm to obtain average-consensus. Sec.V considers the numerical implementation of the four algorithms together with the two comparison algorithms Gossip and ARIS. Resource costs are given in Sec.V-A, and numerical simulations are presented in Sec.V-B. A summary of the results and suggested future work is provided Sec.VI. The Appendix VII presents the proofs of all theorems, the two comparison algorithms, four conjectures, the resource cost derivations, and an important example.

II. The Average-Consensus Problem and Algorithm Assumptions

This section formulates the average-consensus problem and lists the assumptions regarding communication between nodes. Sec.II-A below defines the graph theoretic model for consensus formation that will be subsequently analyzed, and Sec.II-B defines the class of the distributed algorithms we consider. Sec.II-C details the remaining assumptions that will be made on the node communication and computational abilities, and also explains the technique by which the four proposed algorithms will obtain average-consensus.

A. Problem Formulation

Let $t \geq 0$ denote time (the results of this paper assume a continuous time framework; any discrete time index is a special case of this framework). At initial time $t = 0$, consider a finite set of arbitrarily numbered nodes $\mathcal{V} = \{1, \ldots, n\}$ and a set of d-dimensional vectors $\{s_i(0) \in \mathbb{R}^d \ : \ i \in \mathcal{V}\}$. The set $\{s_i(0) \in \mathbb{R}^d \ : \ i \in \mathcal{V}\}$ is referred to as the set of "initial consensus vectors". Suppose each node can locally store a "knowledge set" $\mathcal{K}_i(t)$ that consists of a group of scalars each with a distinct meaning.

- (A1): (knowledge set assumption) At any time $t \geq 0$, each node $i \in \mathcal{V}$ is equipped with a device that can update and store a "knowledge set" $\mathcal{K}_i(t)$. For each $i \in \mathcal{V}$, the knowledge set $\mathcal{K}_i(t)$ may have a time-varying cardinality.

Next we assume that a set $S^{ij}(t_0^{ij}, t_1^{ij})$ can be transmitted from node j at a time denoted $t_0^{ij} \geq 0$, and received at node i at a time denoted as t_1^{ij}, where due to causality $t_1^{ij} \geq t_0^{ij}$. We refer to $S^{ij}(t_0^{ij}, t_1^{ij})$ as a "signal", or "signal set".

- (A2): (signal set assumption) At any time $t_0^{ij} \geq 0$, each node $j \in \mathcal{V}$ has the ability to transmit a "signal set" $S^{ij}(t_0^{ij}, t_1^{ij}) \subseteq \mathcal{K}_j(t_0^{ij})$ that will be received at some node $i \in \mathcal{V}$ at time $t_1^{ij} \geq t_0^{ij}$.

As our final condition, we assume that at $t = 0$ each node $i \in \mathcal{V}$ will "know" its unique node identifier value i, the network size n, and the respective initial consensus vector $s_i(0)$. This is formalized as,

- (A3): At time $t = 0$, the knowledge set $\mathcal{K}_i(0)$ of each node $i \in \mathcal{V}$ satisfies $\mathcal{K}_i(0) \supseteq \{i, n, s_i(0)\}$.

Definition 2.1: Under (A1)-(A3), the average-consensus problem is solved at some time instant t if and only if (iff) the *average* of the initial consensus vectors,

$$\bar{s}(0) = \frac{1}{n} \sum_{i=1}^{n} s_i(0) \ , \tag{1}$$

is contained in the knowledge set $\mathcal{K}_i(t)$ of all nodes $i \in \mathcal{V}$. We say that a specific node i has obtained average-consensus at time t iff $\bar{s}(0) \in \mathcal{K}_i(t)$. □

The four algorithms analyzed in this paper can be adapted so that the size of the network n and unique node identifier $i \in \mathcal{V}$ can be computed during the averaging process, and thus they need not be initially known at any node. For simplicity, however, it will be assumed that the network size n and unique node identifier i are initially known at each node $i \in \mathcal{V}$.

B. The Distributed Algorithm and Local Consensus Estimates

To provide a unifying theme for the algorithms discussed in this paper, we first define the class of distributed algorithms that will be considered for consensus formation. Given the assumptions $(A1) - (A2)$, we define a "distributed algorithm" in terms of its "knowledge set updating rule" $f_{\mathcal{K}}\{\cdot\}$ together with a "signal specification rule" $f_S\{\cdot\}$. The knowledge set updating rule $f_{\mathcal{K}}\{\cdot\}$ defines the effect that a set of signals has on the knowledge set $\mathcal{K}_i(t_1^{ij})$ at the receiving node i, whereas the signal specification rule $f_S\{\cdot\}$ defines the elements contained in the signal $S^{ij}(t_0^{ij}, t_1^{ij})$ given as a function of the knowledge set $\mathcal{K}_j(t_0^{ij})$ at the transmitting node j. If we assume that upon reception of a set of signals $\Delta t > 0$ time elapses before the knowledge set is fully updated, then the class of distributed algorithms we consider may be defined as follows.

Class of Distributed Algorithms under (A1)-(A2):

Knowledge Set Updating Rule: $\qquad f_{\mathcal{K}} \ : \ \mathcal{K}_i(t_1^{ij}) \bigcup_{t_0^{ij} \leq t_1^{ij} \ : \ j \in \mathcal{V}} S^{ij}(t_0^{ij}, t_1^{ij}) \longmapsto \mathcal{K}_i(t_1^{ij} + \Delta t) \qquad (2)$

Signal Specification Rule: $\qquad f_S \ : \ \mathcal{K}_j(t_0^{ij}) \longmapsto S^{ij}(t_0^{ij}, t_1^{ij}) \qquad (3)$

The algorithms in this paper will assume that every knowledge set $\mathcal{K}_i(t)$ contains a local "consensus estimate" $\hat{s}_i(t) \in \mathbb{R}^d$ that represents the "belief" of node i in regard to the average-consensus $\bar{s}(0)$ defined in (1),

$$\mathcal{K}_i(t) \supseteq \{\hat{s}_i(t)\} \ , \ \forall \ i \in \mathcal{V} \ , \ \forall \ t \geq 0 \ . \tag{4}$$

Notice that using a local consensus estimate $\hat{s}_i(t)$ is necessary for any algorithm seeking to solve the average-consensus problem; if there is no local consensus estimate then there is no means by which the knowledge set of any node can contain $\bar{s}(0)$ and hence the problem stated in Definition 2.1 is ill-posed.

In contrast from any known past consensus algorithm, the four proposed algorithms here will also assume that at any time $t \geq 0$ the knowledge set $\mathcal{K}_i(t)$ of each node $i \in \mathcal{V}$ will contain a local "normal consensus estimate" $\hat{v}_i(t) \in \mathbb{R}^n$,

$$\mathcal{K}_i(t) \supseteq \{\hat{v}_i(t)\} \ , \ \forall \ i \in \mathcal{V} \ , \ \forall \ t \geq 0 \ . \tag{5}$$

Combining $(A3)$, (4), and (5) we then have the initial knowledge set for each of the four proposed algorithms,

- $(A4)$: $\mathcal{K}_i(0) = \{i, n, s_i(0), \hat{s}_i(0), \hat{v}_i(0)\} \ , \ \forall \ i \in \mathcal{V} \ .$

Define the "error" of the consensus estimate $\hat{s}_i(t)$ as follows,

$$E_{\bar{s}(0)}\big(\hat{s}_i(t)\big) = |\sqrt{\big(\hat{s}_i(t) - \bar{s}(0)\big)^2}| \ ,$$

where $\bar{s}(0)$ is defined in (1). Denote the "network consensus error" $\sum_{i=1}^n E_{\bar{s}(0)}\big(\hat{s}_i(t)\big)$. We conclude this section with the following definition of what constitutes the solution to the average-consensus problem.

Definition 2.2: The average-consensus problem is solved at time t iff the network consensus error is zero, that is,

$$\sum_{i=1}^n E_{\bar{s}(0)}\big(\hat{s}_i(t)\big) = 0 \ . \tag{6}$$

$\square$

C. Node Communication and Update Assumptions

Below we detail the node communication and update assumptions that will be used by each of the four proposed algorithms. The final update condition we propose will explain the technique by which all four algorithms achieve average-consensus.

For any $t \geq 0$, define $t(+)$ as the right-hand limit of t, that is $t(+) = \lim_{T \to t^+} T$. To construct a suitable average-consensus algorithm we assume the following conditions on the node communication and knowledge set updates:

- (A5): at no time does any node $j \in \mathcal{V}$ have the ability to know *when* the signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is transmitted, *when* the signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received, or *what* node $i \in \mathcal{V}$ will receive the signal $S^{ij}(t_0^{ij}, t_1^{ij})$.
- (A6): at no time does any node $i \in \mathcal{V}$ have the ability to control *when* the signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received, *what* node $j \in \mathcal{V}$ transmitted the signal $S^{ij}(t_0^{ij}, t_1^{ij})$, or *when* the signal $S^{ij}(t_0^{ij}, t_1^{ij})$ was transmitted.
- (A7): each knowledge set satisfies $\mathcal{K}_i(t(+)) = \mathcal{K}_i(t)$ at any time $t \geq 0$ that node i does not receive a signal (recall that $t(+)$ denotes the right-hand limit of t).
- (A8): when a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received, the knowledge set $\mathcal{K}_i(t_1^{ij})$ of the receiving node is updated in an arbitrarily small amount of time.
- (A9): at most one signal can be received and processed by a given node at any given time instant.

Note that (A5)-(A6) imply the communication process is *a priori* unknown at every node. The assumption (A7) implies that the node knowledge set $\mathcal{K}_i(t)$ can only change if a signal is received at node i, and (A8) implies that any signal $S^{ij}(t_0^{ij}, t_1^{ij})$ transmitted from node j is allowed to contain information that has been updated by a signal received at node j at any time preceding t_0^{ij}. Notice that (A8) is realistic since all four of the proposed algorithms will require only a few arithmetic operations in the update process. Assumption (A9) is a technical requirement that simplifies the proofs of convergence. Together (A8) and (A9) imply the knowledge set updating rule $f_{\mathcal{K}}\{\cdot\}$ defined in (2) reduces to,

$$f_{\mathcal{K}} \;:\; \mathcal{K}_i(t_1^{ij}) \cup S^{ij}(t_0^{ij}, t_1^{ij}) \;\longmapsto\; \mathcal{K}_i(t_1^{ij}(+)) \;. \tag{7}$$

Each algorithm we propose can be easily adapted if (A8)-(A9) were relaxed, however this would be at the expense of simplicity in both our framework and analysis.

As our final update condition, we require that the local consensus estimate $\hat{s}_i(t_1^{ij})$ at the receiving node i, as defined above (4), is updated based on the updated normal consensus estimate $\hat{v}_i(t_1^{ij}(+))$ via the relation,

$$\hat{s}_i(t_1^{ij}(+)) = \mathbf{S}\hat{v}_i(t_1^{ij}(+)) \;, \tag{8}$$

where $\mathbf{S} = [s_1(0); s_2(0); \cdots; s_n(0)] \in \mathbb{R}^{d \times n}$. Under (8) it is clear that $\hat{s}_i(t_1^{ij}(+)) = \bar{s}(0)$ if $\hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n$, where $\mathbf{1}_n \in \mathbb{R}^n$ denotes a vector that consists only of unit-valued elements. Thus, in terms of Definition 2.2, under (8) and (5) the average-consensus problem is solved at time t if $\hat{v}_i(t) = \frac{1}{n}\mathbf{1}_n$ for all nodes $i \in \mathcal{V}$. Motivated by this fact, we propose for each of the four algorithms that the normal consensus estimate $\hat{v}_i(t)$ is updated based on the following optimization problem,

$$\begin{aligned} \hat{v}_i(t_1^{ij}(+)) = \quad & \arg_{\tilde{v}} \; \min \; \left(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\right)^2, \\ & \text{s.t. (8) holds, given } \mathcal{K}_i(t_1^{ij}) \bigcup S^{ij}(t_0^{ij}, t_1^{ij}). \end{aligned} \tag{9}$$

Note that from (4) and (5) each node $i \in \mathcal{V}$ will know it has obtained the true average-consensus value $\hat{s}_i(t) = \bar{s}(0)$ when the local normal consensus estimate $\hat{v}_i(t)$ satisfies the condition $\hat{v}_i(t) = \frac{1}{n}\mathbf{1}_n$. In the next section we will define the knowledge set updating rule $f_{\mathcal{K}}\{\cdot\}$ and signal specification rule $f_S\{\cdot\}$ for each of the four algorithms. We shall find that for each algorithm the update problem (9) reduces to a least-squares optimization and a closed-form expression can be obtained for updated normal consensus estimate $\hat{v}_i(t_1^{ij}(+))$.

III. DISTRIBUTED AVERAGE-CONSENSUS ALGORITHMS

With the above definitions, we are now ready to describe the four distributed algorithms that achieve average-consensus. The details of the two comparison algorithms can be found in Sec.VII-B of Appendix VII. All six algorithms (the four presented below and the two comparison algorithms in Sec.VII-B) are defined using the abstract definition of the "distributed algorithm" given in Sec.II, that is (3), (7). This section sets the stage for the convergence theorems presented in Sec.IV.

A. Algorithm 1: Bench-Mark (BM)

The BM algorithm obtains average-consensus trivially; we propose it formally because the remaining three algorithms are specific reductions of it, and also because it requires communication conditions that will be used in our main result. The BM algorithm implies that the initial consensus vectors $\{s_i(0) \; : \; i \in \mathcal{V}\}$ are essentially flooded through-out the network. In [4], [16] the general methodology of the BM algorithm is discussed wherein it is referred to as "flooding". For completeness of our results, we show in Theorems 4.2,4.3 that regardless of the communication pattern between nodes, there exists no distributed algorithm $(2), (3)$ that can obtain average-consensus before the BM algorithm. This is why we have named it the "bench-mark" algorithm.

Let $\delta[\cdot]$ denote the Kronecker delta function applied element-wise, and e_i denote the ith standard unit vector in $\mathbb{R}^n$. The BM signal specification (3) and knowledge set update (7) are respectively defined as (10) and (11) below.

Algorithm 1: Bench-Mark (BM)

Signal Specification:
$$S^{ij}(t_0^{ij}, t_1^{ij}) = \begin{cases} \mathcal{K}_j(t_0^{ij}) \setminus \{j, n, \hat{s}_j(t_0^{ij})\} & \text{if } \hat{v}_j(t_0^{ij}) \neq \frac{1}{n}\mathbf{1}_n \\ \{\hat{v}_j(t_0^{ij}), \hat{s}_j(t_0^{ij})\} & \text{if } \hat{v}_j(t_0^{ij}) = \frac{1}{n}\mathbf{1}_n \end{cases} \tag{10}$$

Knowledge Set Update:

$$v^{ij}(t_1^{ij}, t_0^{ij}) \equiv \mathbf{1}_n - \delta[\hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij})]$$

$$\mathcal{K}_i(t_1^{ij}(+)) = \begin{cases} \{i, n, \hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+)), s_\ell(0) \; \forall \; \ell \text{ s.t. } v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\} & \text{if } \hat{v}_i(t_1^{ij}(+)) \neq \frac{1}{n}\mathbf{1}_n \\ \{\hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+))\} & \text{if } \hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n \end{cases} \tag{11}$$

Normal Consensus Estimate Update:
$$\hat{v}_i(t_1^{ij}(+)) = \frac{1}{n} v^{ij}(t_1^{ij}, t_0^{ij}) \tag{12}$$

Consensus Estimate Update:
$$\hat{s}_i(t_1^{ij}(+)) = \begin{cases} \sum_{\ell=1}^{n} \hat{v}_{i\ell}(t_1^{ij}(+)) s_\ell(0) & \text{if } \hat{v}_i(t_1^{ij}(+)) \neq \frac{1}{n}\mathbf{1}_n \\ \bar{s}(0) & \text{if } \hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n \end{cases} \tag{13}$$

Estimate Initialization:
$$\hat{v}_i(0) = \frac{1}{n} e_i \; , \; \hat{s}_i(0) = \frac{1}{n} s_i(0) \; . \tag{14}$$

In Lemma 7.3 and Lemma 7.2 of Appendix VII we will prove that (12) is the unique solution to (9) under (10) and (11), and that (14) is the unique solution to (9) under the initial knowledge set $(A4)$. The update (13) follows immediately from (12) and the relation (8). Notice that the BM algorithm updates (11), (12) together with the signal specification (10) imply $\bar{s}(0) \in \mathcal{K}_i(t_1^{ij}(t_1^{ij}(+)))$ iff $\hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n$, and likewise $s_\ell(0) \in \mathcal{K}_i(t_1^{ij}(+))$ iff $\hat{v}_{i\ell}(t_1^{ij}(+)) = \frac{1}{n}$ and $\hat{v}_i(t_1^{ij}(+)) \neq \frac{1}{n}\mathbf{1}_n$. It thus follows that (10), (11) and (12) imply that the consensus estimate update $\hat{s}_i(t_1^{ij}(+))$ defined in (13) can be locally computed at node i based only on $\mathcal{K}_i(t_1^{ij})$ and the received signal $S^{ij}(t_0^{ij}, t_1^{ij})$.

Besides flooding the initial consensus vectors, the BM algorithm $(10) - (14)$ has an additional feature that is not necessary but is rather practical: if $\hat{v}_j(t) = \frac{1}{n}\mathbf{1}_n$ then all n of the initial consensus vectors are stored in $\mathcal{K}_j(t)$,

when this occurs the consensus estimate $\hat{s}_j(t)$ defined in (13) will equal $\bar{s}(0)$ and thus node j has obtained average-consensus. In this case all of the initial consensus vectors are deleted from $\mathcal{K}_j(t)$, and any signal transmitted from j contains only the consensus estimate $\hat{s}_j(t) = \bar{s}(0)$ and the vector $\hat{v}_j(t) = \frac{1}{n}\mathbf{1}_n$. Upon reception of a signal containing $\hat{v}_j(t) = \frac{1}{n}\mathbf{1}_n$, the receiving node i can delete all of their locally stored initial consensus vectors and set $\hat{s}_i(t) = \hat{s}_j(t) = \bar{s}(0)$. In this way the average-consensus value $\bar{s}(0)$ can be propagated through-out the network without requiring all n initial consensus vectors to be contained in every signal. See Conjecture 7.4 in Sec. VII-A for a conjecture regarding the BM algorithm.

B. Algorithm 2: Distributed Averaging (DA)

The DA algorithm that we now introduce, to the best of our knowledge, is new. To define the DA update procedure, let V^+ denote the pseudo-inverse of an arbitrary matrix V. The DA signal specification (3) and knowledge set update (7) are defined as (15) and (16) below.

<div align="center">

Algorithm 2: Distributed Averaging (DA)

</div>

Signal Specification: $\qquad S^{ij}(t_0^{ij}, t_1^{ij}) = \mathcal{K}_j(t_0^{ij}) \setminus \{j, n, s_j(0)\}$ $\qquad\qquad$ (15)

Knowledge Set Update: $\quad \mathcal{K}_i(t_1^{ij}(+)) = \{i, n, \hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+)), s_i(0)\}$ $\qquad\qquad$ (16)

Normal Consensus Estimate Update: $\hat{v}_i(t_1^{ij}(+)) = V_{(DA)} V_{(DA)}^+ \frac{1}{n}\mathbf{1}_n$, $V_{(DA)} = [\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), \frac{1}{n}e_i]$ $\quad$ (17)

Consensus Estimate Update: $\quad \hat{s}_i(t_1^{ij}(+)) = V_s V_{(DA)}^+ \frac{1}{n}\mathbf{1}_n$, $V_s = [\hat{s}_i(t_1^{ij}), \hat{s}_j(t_0^{ij}), \frac{1}{n}s_i(0)]$ $\qquad$ (18)

Estimate Initialization: $\qquad \hat{v}_i(0) = \frac{1}{n}e_i$, $\hat{s}_i(0) = \frac{1}{n}s_i(0)$. $\qquad\qquad\qquad$ (19)

The Lemma 7.2 and Lemma 7.5 in Appendix VII, respectively, prove that (19) is the unique solution of (9) under $(A4)$, and that (17) is the unique solution to (9) under (15) and (16). Based on the normal consensus update (17) the relation (8) reduces to (18). From (18) it is clear that the DA consensus estimate update $\hat{s}_i(t_1^{ij}(+))$ can be locally computed at node i based only on $\mathcal{K}_i(t_1^{ij})$ and the received signal $S^{ij}(t_0^{ij}, t_1^{ij})$. Notice that $V_{(DA)}$ is a $n \times 3$ matrix, thus the pseudo-inverse in (17) will have an immediate closed-form expression, see (70), (72) and (73) in Lemma 7.9. Also note that under the DA algorithm every signal contains only the local consensus estimate $\hat{s}_j(t)$ together with the local normal consensus estimate $\hat{v}_j(t)$.

C. Algorithm 3: One-Hop (OH)

Under the OH algorithm each signal $S^{ij}(t_0^{ij}, t_1^{ij})$ will either contain the local initial consensus vector $s_j(0)$ and transmitting node identity j, or the average-consensus value $\bar{s}(0)$ and a scalar 0 to indicate that the transmitted vector is the true average-consensus value. For this reason the conditions for average-consensus under the OH

algorithm are relatively straight-forward to derive (see Theorem 4.7 in Sec.IV and the proof in Appendix VII). The OH algorithm signal specification and knowledge set update are respectively defined by (20) and (21) below.

Algorithm 3: One-Hop (OH)

$$\text{Signal Specification:} \quad S^{ij}(t_0^{ij}, t_1^{ij}) = \begin{cases} \{j, s_j(0)\} & \text{if } \hat{v}_j(t_0^{ij}) \neq \frac{1}{n}\mathbf{1}_n \\ \{0, \hat{s}_j(t_0^{ij})\} & \text{if } \hat{v}_j(t_0^{ij}) = \frac{1}{n}\mathbf{1}_n \end{cases} \quad (20)$$

$$\text{Knowledge Set Update:} \, \mathcal{K}_i(t_1^{ij}(+)) = \begin{cases} \{i, n, \hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+)), s_i(0)\} & \text{if } \hat{v}_i(t_1^{ij}(+)) \neq \frac{1}{n}\mathbf{1}_n \\ \{\hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+))\} & \text{if } \hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n \end{cases} \quad (21)$$

Normal Consensus Estimate Update:

$$v^{ij}(t_1^{ij}, t_0^{ij}) \equiv \mathbf{1}_n - \delta[\hat{v}_i(t_1^{ij}) + e_j]$$

$$\hat{v}_i(t_1^{ij}(+)) = \begin{cases} \frac{1}{n} v^{ij}(t_1^{ij}, t_0^{ij}) & \text{if } 0 \neq S_1^{ij}(t_0^{ij}, t_1^{ij}) \\ \frac{1}{n}\mathbf{1}_n & \text{if } 0 = S_1^{ij}(t_0^{ij}, t_1^{ij}) \end{cases} \quad (22)$$

$$\text{Consensus Estimate Update:} \quad \hat{s}_i(t_1^{ij}(+)) = \begin{cases} \hat{s}_i(t_1^{ij}) + \left(\frac{1}{n} - \hat{v}_{ij}(t_1^{ij})\right)s_j(0) & \text{if } 0 \neq S_1^{ij}(t_0^{ij}, t_1^{ij}) \\ \bar{s}(0) & \text{if } 0 = S_1^{ij}(t_0^{ij}, t_1^{ij}) \end{cases} \quad (23)$$

$$\text{Estimate Initialization:} \quad \hat{v}_i(0) = \frac{1}{n}e_i \,, \,\, \hat{s}_i(0) = \frac{1}{n}s_i(0) \,. \quad (24)$$

The Lemma 7.2 and Lemma 7.20 in Appendix VII respectively prove that (24) is the unique solution to (9) under $(A4)$, and that (22) is the unique solution to (9) under (20) and (21). Notice that (22) implies the relation (8) reduces to (23). Given (20), (21) and (22) it follows that the OH consensus estimate update $\hat{s}_i(t_1^{ij}(+))$ defined in (23) can be locally computed at node i based only on $\mathcal{K}_i(t_1^{ij})$ and the received signal $S^{ij}(t_0^{ij}, t_1^{ij})$.

D. Algorithm 4: Discretized Distributed-Averaging (DDA)

For the DDA algorithm let the discrete set of vectors $\mathbb{R}^n_{0, \frac{1}{n}}$ be defined,

$$\mathbb{R}^n_{0, \frac{1}{n}} = \{v \in \mathbb{R}^n \, : \, v_\ell \in \{0, 1/n\} \,, \,\, \forall \, \ell = 1, 2, \ldots, n \, \}. \quad (25)$$

The discretized version of (9) that we consider under the DDA algorithm is,

$$\hat{v}_i(t_1^{ij}(+)) = \quad \arg_{\tilde{v}} \, \min \left(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\right)^2,$$
$$\text{s.t. (8) holds and } \tilde{v} \in \mathbb{R}^n_{0, \frac{1}{n}}, \text{ given } \mathcal{K}_i(t_1^{ij}) \bigcup S^{ij}(t_0^{ij}, t_1^{ij}). \quad (26)$$

To define the DDA normal consensus update it is convenient to denote $v^{-i} \in \mathbb{R}^{n-1}$ as the vector $v \in \mathbb{R}^n$ with element v_i deleted. The DDA signal specification and knowledge set update are defined below,

---

Algorithm 4: Discretized Distributed-Averaging (DDA)

Signal Specification: $\quad S^{ij}(t_0^{ij}, t_1^{ij}) = \mathcal{K}_j(t_0^{ij}) \setminus \{j, n, s_j(0)\}$ (27)

Knowledge Set Update: $\quad \mathcal{K}_i(t_1^{ij}(+)) = \{i, n, \hat{v}_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+)), s_i(0)\}$ (28)

Normal Consensus Estimate Update:

$$\hat{v}_i(t_1^{ij}(+)) = \hat{a}\hat{v}_i(t_1^{ij}) + \hat{b}\hat{v}_j(t_0^{ij}) + \hat{c}e_i$$ (29)

$\quad (\hat{a}, \hat{b}, \hat{c}) = \left(\ 1,\ 1,\ -\hat{v}_{ji}(t_0^{ij})\ \right) \qquad$ if $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) = 0$

$\quad (\hat{a}, \hat{b}, \hat{c}) = \left(\ 0,\ 1,\ \frac{1}{n} - \hat{v}_{ji}(t_0^{ij})\ \right) \qquad$ if $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$, $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$ (30)

$\quad (\hat{a}, \hat{b}, \hat{c}) = \left(\ 1,\ 0,\ 0\ \right) \qquad$ if $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$, $\hat{v}_i^{-i}(t_1^{ij})^2 \geq \hat{v}_j^{-i}(t_0^{ij})^2$

Consensus Estimate Update: $\quad \hat{s}_i(t_1^{ij}(+)) = \hat{a}\hat{s}_i(t_1^{ij}) + \hat{b}\hat{s}_j(t_0^{ij}) + \hat{c}s_i(0)$ (31)

Estimate Initialization: $\quad \hat{v}_i(0) = \frac{1}{n}e_i$, $\hat{s}_i(0) = \frac{1}{n}s_i(0)$. (32)

Notice that the DDA signal specification (27) and knowledge set update (28) are identical to those of the DA algorithm. Also notice that under (29) the relation (8) implies (31). The Lemma 7.2 and Lemma 7.27 in Appendix VII respectively prove that (32) is the unique solution to (26) under $(A4)$, and that $(29) - (30)$ is a global solution to (26) under (27) and (28). From (27) and (28) it is clear that the DDA consensus estimate update $\hat{s}_i(t_1^{ij}(+))$ defined in (31) can be locally computed at node i based only on $\mathcal{K}_i(t_1^{ij})$ and the received signal $S^{ij}(t_0^{ij}, t_1^{ij})$.

Summary: We have defined four algorithms that in Sec.IV will be shown to achieve average-consensus. The algorithms were derived as special cases of the distributed algorithm $(3), (7)$, where (7) is a special case of (2). Computationally, each algorithm requires only a few elementary arithmetic operations. In Sec.V-A we discuss the storage and communication costs of these four algorithms.

IV. MAIN RESULTS: ANALYSIS OF THE AVERAGE-CONSENSUS ALGORITHMS

This section proves that Algorithms $1-4$ described in Sec.III achieve average-consensus under different communication conditions, we state this below in five theorems. Under assumptions (A1)-(A9) listed in Sec.II-A $-$ II-B, these theorems provide necessary and sufficient conditions on the communication between nodes for average-consensus formation under Algorithms $1-4$ (recall Definition 2.2 defines the solution to the average-consensus problem). The main implication of these results is that, with or without flooding, an average-consensus can be achieved in the presence of arbitrary communication delays and link failures, provided only that there is a uni-directional connectivity condition among nodes.

Each theorem below will assume a certain condition on the communication among nodes. To specify these conditions we require the following two definitions. For any $t_1 \geq t_0 \geq 0$, an arbitrary "communication sequence"

$C_{(t_0,t_1)}$ is defined as the set of all signals transmitted after time t_0 and received before time t_1, that is,

$$C_{[t_0,t_1]} = \{S^{i_1 j_1}, S^{i_2 j_2}, S^{i_3 j_3}, \ldots\}$$

where we have omitted the time indices but it is understood that the transmission time $t_0^{i_\ell j_\ell}$ and reception time $t_1^{i_\ell j_\ell}$ of each signal $S^{i_\ell j_\ell}$ belong to the interval $[t_0, t_1]$. Recall that a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ as stated in (A2) implies that a well-defined subset of $\mathcal{K}_j(t_0^{ij})$ leaves node j at time t_0^{ij} and is received at node i at time t_1^{ij} (the specific subset depends on the algorithm considered).

Next we define the notion of a "communication path". Intuitively, a communication path from node j to i implies that node j transmits a signal received by some node ℓ_1, and *then* node ℓ_1 sends a signal received by some node ℓ_2, and *then* node ℓ_2 sends a signal received by some node ℓ_3, and so on, until node i receives a signal from node $\ell_{k(ij)}$. Technically, we say that $C_{[t_0,t_1]}$ contains a "communication path" from node j to node i iff $C_{[t_0,t_1]}$ contains a sub-sequence $C_{[t_0(ij),t_1(ij)]}^{ij}$ with the following connectivity property,

$$C_{[t_0(ij),t_1(ij)]}^{ij} \supseteq \{S^{\ell_1 j}, S^{\ell_2 \ell_1}, S^{\ell_3 \ell_2}, \\ \ldots, S^{\ell_{k(ij)} \ell_{k(ij)-1}}, S^{i \ell_{k(ij)}}\} \tag{33}$$

where again we have omitted the time indices but it is understood that the transmission time $t_0^{\ell_{q+1} \ell_q}$ of each signal $S^{\ell_{q+1} \ell_q}$ occurs *after* the reception time $t_1^{\ell_q \ell_{q-1}}$ of the preceding signal $S^{\ell_q \ell_{q-1}}$. Note that the communication path $C_{[t_0(ij),t_1(ij)]}^{ij}$ has a cardinality $|C_{[t_0(ij),t_1(ij)]}^{ij}| \geq k(ij) + 1$.

A. Necessary and Sufficient Conditions for Average-Consensus

We are now ready to present the main convergence theorems for Algorithms $1 - 4$.

1) Algorithm 1 (BM): To prove convergence of the BM algorithm, consider the following communication condition. Let us denote $\mathcal{V}_{-i} = \mathcal{V} \setminus \{i\}$ for an arbitrary node $i \in \mathcal{V}$.

Definition 4.1: (S$\mathcal{V}$SC) A communication sequence $C_{[0,t_1]}$ is "singly $\mathcal{V}$-strongly connected" (S$\mathcal{V}$SC) iff there exists a communication path from each node $i \in \mathcal{V}$ to every node $j \in \mathcal{V}_{-i}$. □
We will let "$C_{[t_0,t_1]} \in$ S$\mathcal{V}$SC" denote that a sequence $C_{[t_0,t_1]}$ is S$\mathcal{V}$SC. The definition (33) implies that $C_{[t_0,t_1]} \in$ S$\mathcal{V}$SC iff,

$$C_{[t_0(ij),t_1(ij)]}^{ij} \subset C_{[0,t_1]} , \quad \forall j \in \mathcal{V}_{-i} , \quad \forall i \in \mathcal{V} . \tag{34}$$

The following theorem establishes the sufficient communication conditions for the BM algorithm $(10) - (14)$.

Theorem 4.2: (BM Sufficient Conditions) Consider Algorithm 1, namely the BM algorithm $(10) - (14)$. Then the average-consensus (6) is achieved at time $t = t_1(+)$ for any communication sequence $C_{[0,t_1]}$ satisfying the S$\mathcal{V}$SC condition (34).

Proof: See Appendix VII. ∎

The BM algorithm implies a "flooding" of the initial consensus vectors $\{s_i(0) : i \in \mathcal{V}\}$ through-out the network. There is no other algorithm in the literature that does not use a protocol "equivalent" to the flooding technique and

still guarantees average-consensus for every communication sequence $C_{[0,t_1]}$ satisfying (34), we state this formally as Conjecture 7.4 in Sec.VII-A. Also related to this is Conjecture 7.38 in Sec.VII-B of Appendix VII.

When combined with Theorem 4.2, the following theorem implies that the communication condition *sufficient* for average-consensus under the BM algorithm is the exact same communication condition that is *necessary* for average-consensus under any algorithm. This is why the BM algorithm can be said to possess optimal convergence properties.

Theorem 4.3: (BM,DA,OH and DDA Necessary Conditions) If a communication sequence $C_{[0,t_1]}$ does not satisfy the S$\mathcal{V}$SC condition (34), then no distributed algorithm (2), (3) can achieve average-consensus (6) at time $t = t_1(+)$.

Proof: See Appendix VII. ■

Although Theorem 4.3 is somewhat obvious, it is also a valid necessary communication condition and will be referred to through-out the paper. Note that Theorem 4.7 below states that the OH algorithm requires even stronger necessary conditions than the S$\mathcal{V}$SC condition (34) implied by Theorem 4.3. However, the DA and DDA algorithms do not; there exist many S$\mathcal{V}$SC sequences under which the DA and DDA algorithms will obtain average-consensus at the same instant as the BM algorithm. An example is the "unit-delay double cycle sequence" defined in (170) of Appendix VII. Together with Theorem 4.3, the example (170) implies that the DA and DDA algorithms possess the weakest possible necessary conditions for average-consensus that any algorithm can have, this is illustrated in Fig.1 below.

2) Algorithm 2 (DA): The convergence of Algorithm 2 depends on the following I$\mathcal{V}$SC condition.

Definition 4.4: (I$\mathcal{V}$SC) A communication sequence $C_{[0,t_1]}$ is "infinitely $\mathcal{V}$-strongly connected" (I$\mathcal{V}$SC) iff for each time instant $t \in [0, t_1)$ there exists a finite span of time $T_t \in (0, t_1 - t)$ such that $C_{[t,t+T_t]}$ satisfies the S$\mathcal{V}$SC condition (34). □

It follows that a sequence $C_{[0,t_1]}$ is I$\mathcal{V}$SC iff we can define the infinite set of ordered pairs $\{(t_0^\ell, t_1^\ell) \ : \ \ell \in \mathbb{N}\}$,

$$t_0^0 = 0 \ , \ t_1^\ell = \arg\min_t\{C_{[t_0^\ell,t]} \in \text{S}\mathcal{V}\text{SC}\}$$
$$t_0^\ell = t_1^{\ell-1}(+) \ , \ \ell = 1, 2, \dots.$$

A sequence $C_{[0,t_1]}$ being I$\mathcal{V}$SC is then equivalent to the condition,

$$
\begin{aligned}
C_{[0,t_1]} &= \bigcup_{\ell\in\mathbb{N}} \ C_{[t_0^\ell,t_1^\ell]} \\
C_{[t_0^\ell,t_1^\ell]} &\in \text{S}\mathcal{V}\text{SC} \ , \ \forall \ \ell \in \mathbb{N} \ .
\end{aligned}
\tag{35}
$$

We now proceed to the convergence of the DA algorithm. The following theorem is our main result.

Theorem 4.5: (DA Sufficient Conditions) Consider Algorithm 2, namely the DA algorithm $(15) - (19)$. Then average-consensus (6) is achieved at time $t = t_1(+)$ for any communication sequence $C_{[0,t_1]}$ satisfying the I$\mathcal{V}$SC condition (35).

Proof: See Appendix VII. ■

The above result is interesting because the I$\mathcal{V}$SC condition (35) assumes a weak recurring connectivity between nodes, and also because the resource costs of the DA algorithm are significantly lower than the BM algorithm (see Sec.V-A). Many papers on average-consensus formation such as [17], [3], [10], [26], [21], [25] assume

communication conditions that are special cases of the IVSC condition. See also Conjecture 7.39 in Sec.VII-B of Appendix VII.

3) Algorithm 3 (OH): The following S$\mathcal{V}$CC condition will be shown sufficient for convergence of the OH algorithm.

Definition 4.6: (S$\mathcal{V}$CC) A communication sequence $C_{[0,t_1]}$ is "singly $\mathcal{V}$-completely connected" (S$\mathcal{V}$CC) iff there exists a node $\hat{i} \in \mathcal{V}$ and a time instant $t_{1/2} \in (0, t_1)$ such that,

$$
\begin{aligned}
S^{\hat{i}j}(t_0^{\hat{i}j}, t_1^{\hat{i}j}) &\in C_{[0,t_{1/2}]} \ , \quad \forall\, j \in \mathcal{V}_{-\hat{i}} \\
S^{j\hat{i}}(t_0^{j\hat{i}}, t_1^{j\hat{i}}) &\in C_{(t_{1/2},t_1]} \ , \quad \forall\, j \in \mathcal{V}_{-\hat{i}} \ .
\end{aligned}
\tag{36}
$$

$\square$

The first line in (36) implies that during the interval $[0, t_{1/2}]$ every node $j \in \mathcal{V}_{-\hat{i}}$ will have transmitted a signal that was received by node $\hat{i}$. The second line in (36) implies that during the interval $(t_{1/2}, t_1]$ the node $\hat{i}$ will have transmitted a signal that is received by each node $j \in \mathcal{V}_{-\hat{i}}$. We will let "$C_{[t_0,t_1]} \in$ S$\mathcal{V}$CC" denote that a sequence $C_{[t_0,t_1]}$ is S$\mathcal{V}$CC. Note that the identity of node $\hat{i}$ need not be known by any node. We now consider convergence of the OH algorithm.

Theorem 4.7: (OH Necessary and Sufficient Conditions) Consider Algorithm 3, namely the OH algorithm $(20)-$ (24). Then average-consensus (6) is achieved at time $t = t_1(+)$ iff the communication sequence $C_{[0,t_1]}$ satisfies the following condition:

- (C): for each node i for which there exists a node $j \in \mathcal{V}_{-i}$ such that $S^{ij}(t_0^{ij}, t_1^{ij}) \notin C_{[0,t_1]}$ for all $(t_0^{ij}, t_1^{ij}) \in \mathbb{R}^2$, there exists a communication path $C_{[t_0(i\ell),t_1(i\ell)]}^{i\ell} \in C_{[0,t_1]}$ from at least one node ℓ such that $S^{\ell j}(t_0^{\ell j}, t_1^{\ell j}) \in C_{[0,t_0(i\ell))}$ for all $j \in \mathcal{V}_{-\ell}$.

Proof: See Appendix VII. ∎

Notice that any communication sequence that satisfies the S$\mathcal{V}$CC condition (36) will also satisfy the condition (C) stated in Theorem 4.7; take $\ell = \hat{i}$, $t_0(j\ell) = t_{1/2}(+)$, and $C_{[t_0(j\ell),t_1(j\ell)]}^{j\ell} = S^{j\hat{i}}(t_0^{j\hat{i}}, t_1^{j\hat{i}})$ for all $j \in \mathcal{V}_{-\ell}$. This relation motivates the position of the OH algorithm in the Venn diagrams presented in Fig.1. See also Remark 7.36 in Sec.VII-A. The condition (C) is more general than S$\mathcal{V}$CC, it implies that each node $i \in \mathcal{V}$ will either receive a signal directly from every other node $j \in \mathcal{V}_{-i}$, or have a communication path from some node ℓ after the node ℓ has received a signal directly from every other node $j \in \mathcal{V}_{-\ell}$. We have defined the S$\mathcal{V}$CC condition not only because it is sufficient for average-consensus under the OH algorithm, but also because it is necessary for the definition of the communication condition I$\mathcal{V}$CC described next.

4) Algorithm 4 (DDA): The sufficient conditions for convergence under Algorithm 4 requires the following definition.

Definition 4.8: (I$\mathcal{V}$CC) A communication sequence $C_{[0,t_1]}$ is "infinitely $\mathcal{V}$-completely connected" (I$\mathcal{V}$CC) iff for each time instant $t \in [0, t_1)$ there exists a finite span of time $T_t \in (0, t_1 - t)$ such that $C_{[t,t+T_t]}$ satisfies the S$\mathcal{V}$CC condition (36).

$\square$

It follows that a sequence $C_{[0,t_1]}$ is IVCC iff we can define the infinite set of ordered pairs $\{(t_0^\ell, t_1^\ell) \ : \ \ell \in \mathbb{N}\}$,

$$t_0^0 = 0 \ , \ t_1^\ell = \arg \min_t \{C_{[t_0^\ell, t]} \in S\mathcal{V}CC\}$$
$$t_0^\ell = t_1^{\ell-1}(+) \ , \ \ell = 1, 2, \ldots .$$

A sequence $C_{[0,t_1]}$ being IVCC is then equivalent to the condition,

$$C_{[0,t_1]} = \bigcup_{\ell \in \mathbb{N}} C_{[t_0^\ell, t_1^\ell]}$$
$$C_{[t_0^\ell, t_1^\ell]} \in S\mathcal{V}CC \ , \ \forall \ \ell \in \mathbb{N} \ . \tag{37}$$

We note that the specific node $\hat{i}$ can vary between each $C_{[t_0^\ell, t_1^\ell]}$ sequence, and furthermore we do not assume that any node knows the identity of $\hat{i}$. The following theorem deals with the convergence of Algorithm 4.

Theorem 4.9: (DDA Sufficient Conditions) Consider Algorithm 4, namely the DDA algorithm $(27) - (32)$. Then average-consensus (6) is achieved at some time $t \in (0, t_1)$ for any communication sequence $C_{[0,t_1]}$ satisfying the IVCC condition (37).

Proof: See Appendix VII. ∎

The above result is not quite as interesting as Theorem 4.5, since even though the DDA algorithm is a discretized version of the DA algorithm, the IVCC condition (37) is far stronger than the IVSC condition (35). Also observe that the OH algorithm obtains average-consensus for any $C_{[0,t_1]} \in S\mathcal{V}CC$, whereas the DDA algorithm will obtain average-consensus under the much stronger condition that $C_{[0,t_1]}$ satisfies (37). However, as mentioned above, due to example (170) of Appendix VII the DDA and DA algorithms both can obtain average-consensus under suitable $S\mathcal{V}SC$ sequences, thus they possess much weaker necessary conditions than the OH algorithm. In fact, any communication sequence $C_{[0,t_1]}$ that strictly satisfies (36) implies Algorithms $1 - 4$ all obtain average-consensus at the exact same time instant. This is noteworthy because many past algorithms can only achieve average-consensus asymptotically (e.g. most iterative averaging schemes), in contrast all four of the algorithms considered here can achieve the (finite) bench-mark time for average-consensus under appropriate communication sequences (e.g. any sequence $C_{[0,t_1]}$ that strictly satisfies (36)).

Summary. Theorems $4.2, 4.3, 4.5, 4.7, 4.9$ above state necessary and sufficient communication conditions for average-consensus under Algorithms $1 - 4$ given the assumptions (A1)-(A9). Each theorem is associated with one of four connectivity assumptions on the communication sequence $C_{[0,t_1]}$, denoted by $\{$ $S\mathcal{V}SC$, $I\mathcal{V}SC$, $S\mathcal{V}CC$, $I\mathcal{V}CC$ $\}$ defined in (34),(35),(36), and (37). Observe that assumptions $I\mathcal{V}SC$ and $S\mathcal{V}CC$ are sufficient conditions for $S\mathcal{V}SC$. Furthermore, $I\mathcal{V}CC$ implies that both $I\mathcal{V}SC$ and $S\mathcal{V}CC$ are satisfied, see Fig.1. Notice that each connectivity condition assumes a set of directed signals with an arbitrary delay in the transmission time of each signal. This is significant because, apart from the flooding technique, no other consensus protocol in the current literature can ensure average-consensus in the presence of arbitrary delays in the transmission time of each signal for *a priori* unknown communication sequences. Of course if the communication sequence is known *a priori*, then specific update protocols can always be constructed that guarantee average-consensus at the same instant as the BM algorithm. Current results on average-consensus that do allow communication delays assume the delays have pre-determined upper-bounds, and also require the use of averaging weights that are globally balanced and pre-determined off-line

(see for example [21], [15], [28]). On a related note, besides flooding there appears to be no consensus protocol in the literature that has been proven to guarantee average-consensus for *every* communication sequence $C_{[0,t_1]}$ satisfying any one of the conditions (34),(35),(36), or (37). The majority of past results on average-consensus either require special cases of the I$\mathcal{V}$SC condition, or else can only guarantee approximate average-consensus under the S$\mathcal{V}$SC condition [19], [6], [1]. On the other hand, the two non-trivial algorithms DA and DDA require a set with cardinality upper-bounded bounded by $O(n+d)$ to be communicated and stored at each node, and all previous algorithms besides the flooding and randomized protocols do not possess this drawback.

The Venn diagrams in Fig.1 summarize Theorems $4.2, 4.3, 4.5, 4.7, 4.9$, as well as their relation to the Gossip and ARIS algorithms that will be used as comparisons to the four proposed algorithms. It remains an open problem whether any protocol exists that guarantees average-consensus for all I$\mathcal{V}$SC sequences without requiring a set with cardinality upper bounded by $O(n+d)$ to be stored and communicated at each node.
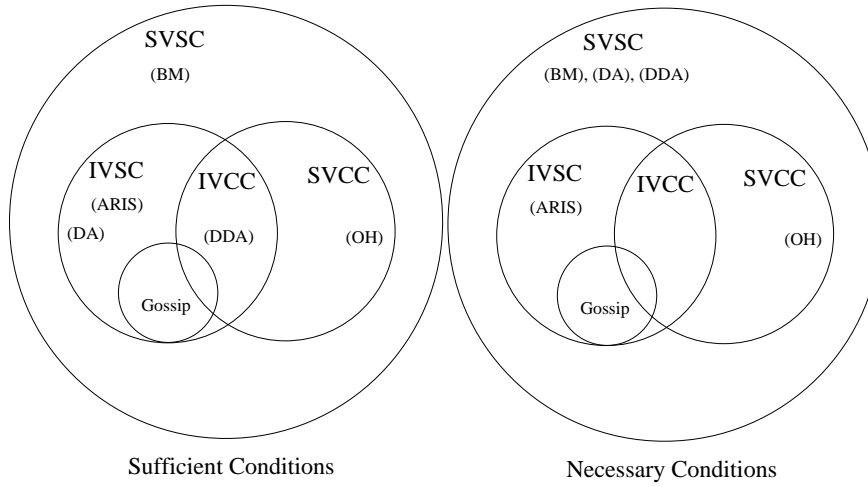

Figure 1. Venn Diagram of Sufficient and Necessary Conditions for Algorithms 1-4 as well as the Comparison Algorithms Gossip and ARIS to Achieve Average-Consensus. The condition (C) in Theorem 4.7 is omitted for simplicity of presentation.

V. Numerical Implementation and Examples

This section presents the communication and storage costs of the various algorithms proposed above. In Sec.V-A below we define the resource costs for the six algorithms compared in this paper (four proposed in Sec.III and two others in Sec.VII-B). In Sec.V-B we present numerical simulations of the six algorithms under various randomized full network graphs when assuming instantaneous communication.

A. Algorithm Resource Costs

Each of the four average-consensus algorithms as well as the two comparison algorithms presented in Sec.VII-B require that the knowledge set $\mathcal{K}_i(t)$ and signal set $S^{ij}(t_0^{ij}, t_1^{ij})$ are respectively defined by a set of scalars, where each scalar has a particular meaning. We are thus motivated to quantify the "resource cost" of each algorithm

Algorithm	BM	DA	OH	DDA	Gossip	ARIS
$\min(\phi_i(t))$	$4d+5$	$4d+6$	$4d+5$	$4d+5$	$2d$	$7+2(r+2)d$
$\max(\phi_i(t))$	$2nd+4+\lfloor\frac{n}{2}\rfloor$	$4d+2n+4$	$4d+4+\lfloor\frac{n}{2}\rfloor$	$4d+4+\lfloor\frac{n}{2}\rfloor$	$2d$	$\lfloor\frac{n}{2}\rfloor+6+2(r+2)d$
$\min(\rho_{ij}(t))$	$2d+1$	$2d+1$	$2d+2$	$2d+1$	$2d$	$3+2(r+1)d$
$\max(\rho_{ij}(t))$	$2(n-1)d+\lfloor\frac{n}{2}\rfloor$	$2d+2n$	$2d+2$	$2d+\lfloor\frac{n}{2}\rfloor$	$2d$	$\lfloor\frac{n}{2}\rfloor+2+2(r+1)d$
$\min(\omega_i(t))$	$6d+6$	$6d+8$	$6d+7$	$6d+6$	$4d$	$10+4(r+\frac{3}{2})d$
$\max(\omega_i(t))$	$2(2n-1)d+4+2\lfloor\frac{n}{2}\rfloor$	$6d+4n+4$	$6d+6+\lfloor\frac{n}{2}\rfloor$	$6d+4+2\lfloor\frac{n}{2}\rfloor$	$4d$	$2\lfloor\frac{n}{2}\rfloor+8+4(r+\frac{3}{2})d$
$O(\phi)$	nd	$n+d$	$n+d$	$n+d$	d	$rd+n$
$O(\rho)$	nd	$n+d$	d	$n+d$	d	$rd+n$

Table I

MINIMUM AND MAXIMUM RESOURCE COSTS OBTAINED BY APPLYING EACH ALGORITHM IN SEC.III TO (38), (39), AND (40). HERE d DENOTES THE DIMENSION OF THE INITIAL CONSENSUS VECTORS, n DENOTES THE NUMBER OF NODES, $\rho_{ij}(t)$ DENOTES THE SIGNAL SET DIMENSION, $\phi_i(t)$ THE KNOWLEDGE SET DIMENSION, AND $\omega_i(t)$ IS THE SUM OF $\rho_{ij}(t)$ AND $\phi_i(t)$. WE LET $\lfloor\cdot\rfloor$ DENOTE THE "FLOOR" OPERATION.

in terms of the total number of scalar values that are required to define the two sets $\mathcal{K}_i(t)$ and $S^{ij}(t_0^{ij}, t_1^{ij})$. In particular, for any $t \geq 0$ we define the "storage cost" of an arbitrary node $i \in \mathcal{V}$ as $\phi_i(t)$,

$$\phi_i(t) = \text{the minimum number of scalars}$$
$$\text{required to define the knowledge set } \mathcal{K}_i(t). \tag{38}$$

Likewise we define the "communication cost" $\rho_{ij}(t_0^{ij})$ of an arbitrary signal $S^{ij}(t_0^{ij}, t_1^{ij})$ as follows,

$$\rho_{ij}(t_0^{ij}) = \text{the minimum number of scalars}$$
$$\text{required to define the set } S^{ij}(t_0^{ij}, t_1^{ij}). \tag{39}$$

Next define the total resource cost of node i at time $t \geq 0$ as $\omega_i(t)$,

$$\omega_i(t) = \phi_i(t) + \rho_{ij}(t) . \tag{40}$$

Based on the knowledge set and signal specifications defined in Sec.III, the Table I presents the resource cost computations of each algorithm when using the definitions (38), (39), and (40). The entries of Table I are derived in Sec.VII-C of Appendix VII.

Note that the storage cost $\phi_i(t)$ is defined per node, and the communication cost $\rho_{ij}(t)$ is defined per signal. The total maximum resource costs of the BM algorithm increase on the order $O(nd)$, whereas the total maximum resource costs of the DA and DDA algorithm increase on the order $O(n + d)$. Although the total maximum resource costs of the OH algorithm increase on the order $O(n + d)$, the maximum resource cost of each signal under the OH algorithm increases only on the order $O(d)$. However, the communication conditions necessary for average-consensus under the OH algorithm are much stronger than the conditions necessary under the BM, DA, and DDA algorithms. This disparity makes it difficult to state definitive results regarding the least costly algorithm under general communication sequences. If condition (C) in Theorem 4.7 is known to hold *a priori*, then the OH algorithm may be preferable to the DA and DDA due to the lower communication costs $O(d) \leq O(d + n)$. On

the other hand, under example (170) in Appendix VII the DDA algorithm is preferable to both the DA and OH algorithm, since the former implies larger resource costs and the latter will not obtain average-consensus.

The Gossip algorithm in [3] has total resource costs that increase on the order $O(d)$, however this algorithm requires strictly bi-directional and instantaneous communication, see Sec.VII-B as well as Fig.2 in Sec.V-B. The total resource costs of the ARIS algorithm increase on the order $O(rd+n)$, where r is an ARIS algorithm parameter explained in Sec.VII-B. If $r \geq n$ then ARIS is more costly than the BM algorithm. For this reason, the simulations presented in Sec.V-B assume $r = n$. The RIS algorithm proposed in [19] requires r random variables to be initially generated at each node for each element of the respective local initial consensus vector $s_i(0) \in \mathbb{R}^d$, the RIS algorithm also requires these random variables to be communicated between nodes, thus both the storage and communication costs of the RIS algorithm increase on the order $O(rd)$.

B. Numerical Results

We present here numerical simulations of the four proposed algorithms together with the two comparison algorithms. The algorithm parameters were chosen as $n = 80$ (number of nodes), $d = 1$ (dimension of the initial consensus vectors), $r = n$ (number of ARIS random variables generated per initial consensus element), and $s_i(0) = i, \ \forall \ i \in \mathcal{V}$ (initial consensus vector values). The node communication is assumed to be instantaneous and in discrete-time, using the following randomized protocol:

- at each $k \in \mathbb{N}$, two nodes $(i, j) \in \mathcal{V}^2$ are uniformly chosen at random such that $i \neq j$,
- with probability one, the node i sends a signal to the node j at time k,
- with probability p the node j sends a signal to the node i at time k.

We compare four choices of p, namely $p \in \{1, \frac{1}{2}, \frac{1}{4}, 0\}$. Note that $p = 1$ implies instantaneous bi-directional communication. As p decreases there will be fewer expected signals per time instant, thus we expect that each algorithm will have slower convergence for lower values of p.

Fig.2 shows the network consensus error under each algorithm for each value of $p \in \{1, \frac{1}{2}, \frac{1}{4}, 0\}$. It is clear that the Gossip algorithm only converges to average-consensus for $p = 1$. For $p \in \{\frac{1}{2}, \frac{1}{4}, 0\}$ the Gossip algorithm converges to a consensus that is increasingly distant from the average-consensus. The BM algorithm converges fastest in all simulations, as expected from Theorems 4.2, 4.3. Initially the DA algorithm can be observed to almost match the BM algorithm in all simulations, however as time proceeds for $p = 1$ the Gossip and ARIS algorithm eventually overcome DA, and in $p \in \{\frac{1}{2}, \frac{1}{4}, 0\}$ the ARIS algorithm eventually overcomes DA. However, the resource costs of ARIS in our simulations is greater than even the BM algorithm, thus further research is needed to objectively evaluate the trade-off between resource cost and network consensus error for each algorithm under various communication conditions.

VI. Conclusion

This paper has described and analyzed four distributed algorithms designed to solve the average-consensus problem under general uni-directional connectivity conditions. We derived necessary and sufficient conditions for
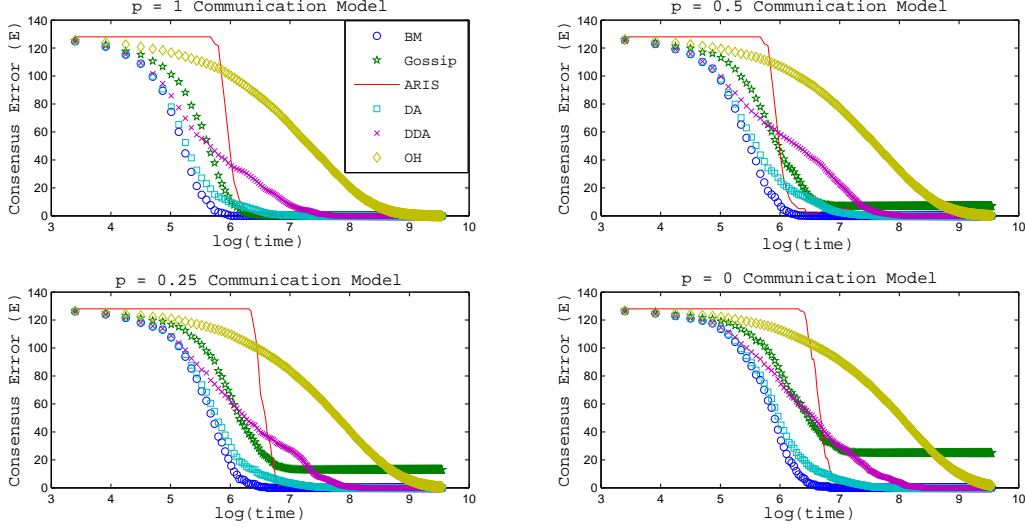
Figure 2. Convergence of Consensus Algorithms for Communication Probabilities $p \in \{1, \frac{1}{2}, \frac{1}{4}, 0\}$ (top left to bottom right). The consensus error E is defined above (6). Observe that the Gossip consensus error vanishes only for $p = 1$, because this is the only model that implies strictly bi-directional communication. All algorithms converge moderately slower as p decreases.

the convergence to average-consensus under each respective algorithm. The conditions for convergence were based on two types of connectivity conditions, namely the singly $\mathcal{V}$-strongly connected sequence (defined in Sec.IV-A1), and the singly $\mathcal{V}$-completely connected sequence (defined in Sec.IV-A3). Both connectivity conditions allow arbitrary delays in the transmission time of each signal, and we did not assume that the sending node knows the identity of the receiving node. The resource costs for each of the algorithms were derived and shown to differ in regard to their order of magnitude with respect to the parameters n (number of nodes) and d (dimension of the initial consensus vectors). Comparisons were made with two known consensus algorithms, referred to as the Gossip algorithm and the adapted randomized information spreading (ARIS) algorithm. Simulations were provided as well as Venn diagrams of the connectivity conditions required for average-consensus under each algorithm. The non-trivial algorithms considered here are relatively advantageous under weak communication conditions if the dimension d of the initial consensus vectors exceeds the network size n. The works [4] and [20] provide two practical examples of when $d \geq n$ might typically be the case, e.g. distributed inference regarding a $\sqrt{n}$ dimensional process or parameter in noise.

The four communication conditions we proposed were deterministic; there were no stochastic properties assumed in regard to the signal process between nodes. However, our framework allowed directed signals as well as arbitrary delays in transmission time, hence *every* causal signal sequence is a special case of our framework. This suggests that the four proposed algorithms can be applied to stochastic communication models for which there is a non-zero probability of consensus under any distributed algorithm, however future work is needed in this direction.

Future work is also needed to obtain a lower bound on the convergence rate to average-consensus for given

characterizations of the communication sequence under each proposed algorithm, as well as improved algorithms designed specifically for particular communication sequences. For an objective evaluation of the various average-consensus algorithms, additional research is needed to compare the evolution of a cost function of the resource cost and network consensus error under a variety of communication sequences. Lastly, an interesting generalization of the average-consensus problem is to allow the initial consensus vectors $s_i(0)$ to vary with time, as discussed for instance by [20], [8], [22], [13], [29], [16]. Applying the algorithms and communication conditions proposed in this work could yield further results with regard to ensuring the average $\bar{s}(t)$ is contained in the knowledge set $\mathcal{K}_i(t)$ of each node $i \in \mathcal{V}$ at some time instants t for given dynamic models of the set $\{s_i(t) \ : \ i \in \mathcal{V}\}$.

As a final note, it is worth mentioning that each of the four proposed algorithms can obtain a consensus on any *linear combination* of the initial consensus vectors $\{s_i(0) \in \mathbb{R}^d \ : \ i \in \mathcal{V}\}$ under the exact same communication conditions as stated in the main results. In other words, suppose a vector $w \in \mathbb{R}^n$ was initially known at each node, then if each algorithm updated the normal consensus estimate based on (9) with $\frac{1}{n}\mathbf{1}_n$ replaced by the vector w, the same conditions stated in Sec.IV will imply the respective algorithms ensure $\hat{s}_i(t_1(+)) = \sum_{i \in \mathcal{V}} s_i(0) w_i$ for all $i \in \mathcal{V}$. The proofs of these results follow by identical arguments to those presented in this work, simply by replacing $\frac{1}{n}\mathbf{1}_n$ by the vector w.

VII. APPENDIX

In Sec.VII-A of this appendix we derive the consensus estimate initialization for each of the four proposed algorithms, as well as the proofs for Theorems 4.2,4.3,4.5,4.7, and 4.9. In Sec.VII-B we define the Gossip algorithm proposed in [3] in terms of the class of distributed algorithms $(3), (7)$, and then we define the ARIS algorithm in these terms as well. In Sec.VII-C we derive the resource costs presented in Table I of Sec.V-A, and in Sec.VII-D we define the "unit-delay double cycle sequence" as an example of a $S\mathcal{V}SC$ sequence that implies the DA, DDA, and BM algorithm all obtain average-consensus at the same instant. We present two conjectures in Sec.VII-A, and two conjectures in Sec.VII-B.

Through-out the appendix we denote the "error" of the normal consensus estimate $\hat{v}_i(t)$ as,

$$E_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(t)) = \left| \sqrt{\left(\hat{v}_i(t) - \frac{1}{n}\mathbf{1}_n\right)^2} \right| . \tag{41}$$

The total reduction in normal consensus squared error resulting from the sequence $C_{[t_0,t_1]}$ is then,

$$\mathbf{E}^2(C_{[t_0,t_1]}) \equiv \sum_{S^{ij}(t_0^{ij},t_1^{ij}) \in C_{[t_0,t_1]}} \mathbf{E}^2\big(S^{ij}(t_0^{ij},t_1^{ij})\big) , \tag{42}$$

where $\mathbf{E}^2\big(S^{ij}(t_0^{ij},t_1^{ij})\big)$ is defined using the normal consensus error $E_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}(t)\big)$ in (41),

$$\mathbf{E}^2\big(S^{ij}(t_0^{ij},t_1^{ij})\big) \equiv E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij})\big) - E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij}(+))\big) . \tag{43}$$

It is convenient to use the following definition.

Definition 7.1: Under (8), (4), and (5), the average-consensus problem is solved at time t iff the "network normal consensus error" $\sum_{i=1}^{n} E_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t)\big)$ is zero, that is,

$$\sum_{i=1}^{n} E_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t)\big) = 0 \ , \tag{44}$$

where $E_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t)\big)$ is defined in (41).

A. Algorithm Convergence Proofs

Lemma 7.2: (Consensus Estimate Initialization) Given the initial knowledge state (A4), the unique solution to both (9) and (26) is,

$$\hat{v}_i(0) = \frac{1}{n}e_i \ , \ \ \hat{s}_i(0) = \frac{1}{n}s_i(0) \ . \tag{45}$$

Proof: Under (A4) the update (9) becomes,

$$\begin{aligned} \hat{v}_i(0) = \ \ &\arg_{\tilde{v}} \ \min \ \big(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\big)^2 \ , \\ &\text{s.t. (8) holds, given } \mathcal{K}_i(0) \supseteq \{i, n, s_i(0)\}. \end{aligned} \tag{46}$$

Observe that the vector e_i can be locally constructed at each node $i \in \mathcal{V}$ based only on the data $\{i, n\}$. Next observe that $s_i(0) = \mathbf{S}e_i$ for all $i \in \mathcal{V}$. Given that $s_i(0) = \mathbf{S}e_i$ is known by node i at $t = 0$, the set of vectors $\hat{v}_i(0)$ for which (8) holds is span$\{e_i\}$, thus (46) becomes,

$$\begin{aligned} \hat{v}_i(0) \ &= \arg_{\tilde{v}} \ \min_{\tilde{v} \in \text{span}\{e_i\}} \ \big(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\big)^2 \ , \\ &= VV^{+}\tfrac{1}{n}\mathbf{1}_n \ \ , \ \ V = [e_i] \\ &= V(V'V)^{-1}V'\tfrac{1}{n}\mathbf{1}_n = \tfrac{1}{n}e_i \ . \end{aligned} \tag{47}$$

If $\hat{v}_i(0) = \frac{1}{n}e_i$ then (8) implies $\hat{s}_i(0) = \frac{1}{n}s_i(0)$. Note that $\frac{1}{n}e_i$ and $\frac{1}{n}s_i(0)$ can both be initially computed at node i given that $\mathcal{K}_i(0) \supseteq \{i, n, s_i(0)\}$. The solution (45) implies $\hat{v}_i(0) \in \mathbb{R}^n_{0,\frac{1}{n}}$ and thus (45) is a feasible solution to (26). It then follows that (45) is also the unique solution to (26) since the feasible space of (26) is contained in the feasible space of (9), and both (9) and (26) minimize the same objective function. ∎

Lemma 7.3: (BM Normal Consensus Estimate Update) Applying (10) and (11) to (9) yields the BM normal consensus estimate update (12).

Proof: If $\hat{v}_i(t_1^{ij}) = \frac{1}{n}\mathbf{1}_n$ then (8) implies $\hat{s}_i(t_1^{ij}) = \bar{s}(0)$ and thus node i need not update its knowledge state regardless of the signal $S^{ij}(t_0^{ij}, t_1^{ij})$. If $\hat{v}_j(t_0^{ij}) = \frac{1}{n}\mathbf{1}_n$ then (10) together with (8) implies,

$$\begin{aligned} S^{ij}(t_0^{ij}, t_1^{ij}) \ &\supseteq \hat{s}_j(t_0^{ij}) \ , \\ &\supseteq \bar{s}(0) \ . \end{aligned} \tag{48}$$

In this case (9) can be re-written,

$$\begin{aligned} \hat{v}_i(t_1^{ij}(+)) = \ \ &\arg_{\tilde{v}} \ \min \ \big(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\big)^2 \ , \\ &\text{s.t. (8) holds, given } \hat{s}_j(t_0^{ij}) = \bar{s}(0). \end{aligned} \tag{49}$$

Since $\hat{s}_j(t_0^{ij}) = \bar{s}(0)$ is known we can let $\hat{s}_i(t_1^{ij}(+)) = \bar{s}(0)$ and thus obtain $\hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n$ as the unique global solution to (49) (note that $\frac{1}{n}\mathbf{1}_n$ can be computed since (A4) and (A7) imply $\mathcal{K}_i(t_1^{ij}) \supseteq \{n\}$). Notice that

this solution coincides with (12). Next assume that $\hat{v}_i(t_1^{ij}) \neq \frac{1}{n}\mathbf{1}_n$ and $\hat{v}_j(t_0^{ij}) \neq \frac{1}{n}\mathbf{1}_n$. Under (10) and (11) we can then re-write (9) as,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) = \quad &\arg_{\tilde{v}} \ \min \ \left(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\right)^2, \\
&\text{s.t. (8) holds, given } s_\ell(0) = \mathbf{S}e_\ell, \\
&\text{for all } \{\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\}.
\end{aligned}
\tag{50}
$$

Given that $s_\ell(0) = \mathbf{S}e_\ell$ for all $\{\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\}$ is known, the set of vectors $\hat{v}_i(t_1^{ij}(+))$ for which (8) holds is $\mathrm{span}\{e_\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\}$, thus (50) becomes,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) \ &= \arg_{\tilde{v}} \ \min_{\tilde{v} \,\in\, \mathrm{span}\{e_\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\}} \ \left(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\right)^2 , \\
&= V_{(BM)}V_{(BM)}^+ \tfrac{1}{n}\mathbf{1}_n , \\
&V_{(BM)} = [e_\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1] .
\end{aligned}
\tag{51}
$$

Let $L = v^{ij}(t_1^{ij}, t_0^{ij})'\mathbf{1}_n$ denote cardinality of the set $\{\ell \ : \ v_\ell^{ij}(t_1^{ij}, t_0^{ij}) = 1\}$. Since the columns of $V_{(BM)}$ are linearly independent, the right-hand side (RHS) of (51) can be computed as,

$$
\begin{aligned}
V_{(BM)}V_{(BM)}^+ \tfrac{1}{n}\mathbf{1}_n \ &= \tfrac{1}{n}V_{(BM)}(V_{(BM)}'V_{(BM)})^{-1}V_{(BM)}'\mathbf{1}_n \\
&= \tfrac{1}{n}V_{(BM)}(I_{L\times L})^{-1}\mathbf{1}_L' , \\
&= \tfrac{1}{n}v^{ij}(t_1^{ij}, t_0^{ij}) ,
\end{aligned}
$$

where $I_{L\times L}$ denotes the identity matrix of dimension L. ∎

Theorem 4.2(BM Network Convergence to Average-Consensus)

Proof: By the BM update (12) when any node receives a signal $S^{ij}(t_0^{ij}, t_1^{ij})$, the receiving node i either receives the desired average-consensus value $\bar{s}(0)$, or receives the initial consensus vector $s_\ell(0)$ of every node that has a communication path to node j within the time span $[0, t_0^{ij}]$. If $C_{[0,t_1]}$ satisfies (34) then every node has a communication path to every other node within the time span $[0, t_1]$, thus (A7) and (12) together imply that at time $t_1(+)$ every node i will compute $\hat{v}_i(t_1(+)) = \frac{1}{n}\mathbf{1}_n$. This implies (44) holds at $t_1(+)$ and hence by Definition (7.1) a network average-consensus is obtained at time $t = t_1(+)$. ∎

Theorem 4.3(BM,DA,OH and DDA Necessary Conditions)

Proof: If $C_{[0,t_1]} \notin S\mathcal{V}SC$ then there exists a node $i \in \mathcal{V}$ that does not have a communication path to some node $j \in \mathcal{V}_{-i}$ within the time span $[0, t_1]$. At time t_1 the node j thus cannot have any knowledge that is contained in $\mathcal{K}_i(t)$ for any $t \leq t_1$, regardless of the knowledge set update rule (2) and signal specification (3). Hence $\hat{s}_j(t_1(+))$ cannot be a function of $s_i(0)$ for an arbitrary initial consensus vector $s_i(0)$. It then follows that no distributed algorithm $(2),(3)$ can imply (6) is satisfied at time $t \leq t_1(+)$ for an arbitrary set of initial consensus vectors $\{s_i(0) \ : \ i \in \mathcal{V}\}$. ∎

We now present a conjecture regarding the BM algorithm. From Theorem 4.2 and $(10) - (14)$ it follows that the BM algorithm implies the following property (P) of each knowledge set $\mathcal{K}_i(t)$:

- (P):

$$C^{ij}_{[t_0(ij),t_1(ij)]} \subset C_{[0,t_1]} \Rightarrow \begin{cases} \mathcal{K}_i(t_1(+)) \supseteq \{s_j(0)\} & \text{, if } \exists \ \ell \text{ s.t. } C^{i\ell}_{[t_0(i\ell),t_1(i\ell)]} \nsubseteq C_{[0,t_1]} \\ \mathcal{K}_i(t_1(+)) \supseteq \{\bar{s}(0)\} & \text{, if } C^{i\ell}_{[t_0(i\ell),t_1(i\ell)]} \subset C_{[0,t_1]} \ \forall \ \ell \in \mathcal{V}_{-i} \end{cases}$$

The condition (P) forms an equivalence class among all algorithms $\mathcal{A}$ defined by (2) and (3). From the second line in (P) it follows that any algorithm $\mathcal{A}$ that satisfies (P) will imply (6) holds at time $t = t_1(+)$ for any communication sequence $C_{[0,t_1]}$ satisfying (34). We now conjecture that if an algorithm $\mathcal{A}$ implies (6) holds at time $t = t_1(+)$ for every communication sequence $C_{[0,t_1]}$ satisfying (34), then algorithm $\mathcal{A}$ must have resource costs at least as great as any algorithm $\mathcal{A}$ that satisfies (P).

Conjecture 7.4: If an algorithm $\mathcal{A}$ guarantees (6) holds at time $t = t_1(+)$ for every communication sequence $C_{[0,t_1]}$ satisfying (34), then the algorithm $\mathcal{A}$ will require that a set with cardinality upper bounded by at least $O(nd)$ can be communicated and stored at each node.

The above conjecture implies that any algorithm $\mathcal{A}$ that satisfies (P) will require that a set with cardinality upper bounded by $O(nd)$ can be communicated and stored at each node, this is why searching for less costly algorithms is of importance. The problem is, less costly algorithms tend to require stronger communication conditions than the BM algorithm, and they also do not guarantee average-consensus is obtained as quickly as the BM algorithm. We note that due to the resource costs associated with the RIS and ARIS algorithms, the Conjecture 7.38 in Sec.VII-B2 does not contradict Conjecture 7.4.

Proof. (Theorem 4.5) Lemmas 7.5 - 7.19. *Overview of Proof.* To prove Theorem 4.5 we initially show in Lemma 7.6 that the update (17) implies each normal consensus estimate $\hat{v}_i(t)$ satisfies the normalization property (54). The Lemma 7.9 proves that each normal consensus estimate $\hat{v}_i(t)$ also satisfies the "zero local error" property $\hat{v}_{ii}(t) = 1/n$. Without the latter, the following lemmas would still imply convergence of all consensus estimates, but the final consensus value would not necessarily equal the average $\bar{s}(0)$ defined in (1). The essence of the convergence proof is that the reduction in error that results from any signal will eventually vanish if $C_{[0,t_1]} \in \mathcal{IVSC}$, see Lemma 7.13. Applying this result to the DA lower bounds on the reduction in error derived in Lemmas 7.10 and 7.11, we can show that each normal consensus vector will necessarily converge to a common vector, see Lemma 7.15. Together with the two technical results derived in Lemmas 7.17 and 7.18, we then combine the triangle inequality and the "zero local error" property to prove that the common vector will approach $\frac{1}{n}\mathbf{1}_n$ in the L^2 norm as time approaches t_1. This implies (44) holds at $t_1(+)$ and hence by Definition (7.1) a network average-consensus is obtained at time $t = t_1(+)$.

Lemma 7.5: (DA Normal Consensus Estimate Update) Applying (15) and (16) to (9) yields the DA normal consensus estimate update $\hat{v}_i(t_1^{ij}(+))$ defined in (17).

Proof: Under (15) and (16) we can re-write (9) as,

$$\hat{v}_i(t_1^{ij}(+)) = \arg_{\tilde{v}} \min \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n \right)^2 ,$$
$$\text{s.t. (8) holds, given } \hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij}), \tag{52}$$
$$\hat{s}_j(t_0^{ij}) = \mathbf{S}\hat{v}_j(t_0^{ij}), \text{ and } s_i(0) = \mathbf{S}e_i.$$

Given that $\hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij}), \hat{s}_j(t_0^{ij}) = \mathbf{S}\hat{v}_j(t_0^{ij})$, and $s_i(0) = \mathbf{S}e_i$ are known, the set of vectors $\hat{v}_i(t_1^{ij}(+))$ for which (8) holds is $\text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$, thus (52) can be re-written as

$$\hat{v}_i(t_1^{ij}(+)) = \arg_{\tilde{v}} \min_{\tilde{v} \in \text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}} \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n \right)^2. \tag{53}$$

The update (17) follows immediately from (53). ∎

Lemma 7.6: (DA Consensus Estimate Normalization) Every normal consensus estimate $\hat{v}_i(t)$ satisfies

$$\hat{v}_i(t)^2 = \frac{1}{n}\hat{v}_i(t)'\mathbf{1}_n , \ \forall \ i \in \mathcal{V} , \ \forall \ t \geq 0 . \tag{54}$$

Proof: Note that $\hat{v}_i(0) = \frac{1}{n}e_i$ satisfies (54) for each $i \in \mathcal{V}$. Next observe that under $(A7)$ the estimate $\hat{v}_i(t)$ will not change unless a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received at node i. If a signal is received then by Lemma 7.5 the estimate $\hat{v}_i(t)$ is updated to the unique solution of (53). Thus to finish the proof it suffices to show that if a vector $v \in \mathbb{R}^n$ does not satisfy (54) then the vector v is not the solution to (53). To prove this we show that if (54) does not hold then the vector w defined,

$$w = v(v'\mathbf{1}_n)/(nv^2) ,$$

will satisfy the inequality

$$\left(w - \frac{1}{n}\mathbf{1}_n \right)^2 < \left(v - \frac{1}{n}\mathbf{1}_n \right)^2 . \tag{55}$$

Notice that since w is contained in $\text{span}(v)$, the inequality (55) implies that v is not the solution to (53). Next observe that if a vector v does not satisfy (54) then,

$$\left(v^2 - \frac{1}{n}v'\mathbf{1}_n \right)^2 > 0 . \tag{56}$$

Expanding (56) yields,

$$(v^2)^2 - 2\frac{1}{n}v'\mathbf{1}_n v^2 + \left(\frac{1}{n}v'\mathbf{1}_n \right)^2 > 0 . \tag{57}$$

Re-arranging (57) then implies (55),

$$(v^2)^2 - 2\frac{1}{n}v'\mathbf{1}_n v^2 > -\left(\frac{1}{n}v'\mathbf{1}_n \right)^2 ,$$
$$v^2 - 2\frac{1}{n}v'\mathbf{1}_n + \frac{1}{n} > \left(\frac{v'\mathbf{1}_n}{nv^2} \right)^2 v^2 - 2\frac{(v'\mathbf{1}_n)^2}{n^2 v^2} + \frac{1}{n} ,$$
$$\left(v - \frac{1}{n}\mathbf{1}_n \right)^2 > \left(v\frac{v'\mathbf{1}_n}{nv^2} - \frac{1}{n}\mathbf{1}_n \right)^2 .$$

∎

Lemma 7.7: (DA Non-Decreasing Normal Consensus Estimate Magnitude) Each magnitude $\hat{v}_i(t)^2$ is a non-decreasing function of $t \geq 0$ for all $i \in \mathcal{V}$.

Proof: Note that under (A7) the estimate $\hat{v}_i(t)$ will not change unless a signal is received at node i. If a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received then the DA update problem (53) implies the update $\hat{v}_i(t_1^{ij}(+))$ must satisfy,

$$\left(\hat{v}_i(t_1^{ij}(+)) - \tfrac{1}{n}\mathbf{1}_n\right)^2 \le \left(w - \tfrac{1}{n}\mathbf{1}_n\right)^2 , \qquad (58)$$
$$\forall\, w \in \text{span } \{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}.$$

Since $\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij})\} \in \text{span } \{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$ the inequality (58) implies,

$$\left(\hat{v}_i(t_1^{ij}(+)) - \tfrac{1}{n}\mathbf{1}_n\right)^2 \le \min\{\left(\hat{v}_i(t_1^{ij}) - \tfrac{1}{n}\mathbf{1}_n\right)^2 , \qquad (59)$$
$$\left(\hat{v}_j(t_0^{ij}) - \tfrac{1}{n}\mathbf{1}_n\right)^2\}.$$

Next observe that if a vector $v \in \mathbb{R}^n$ satisfies (54), then,

$$\left(v - \tfrac{1}{n}\mathbf{1}_n\right)^2 \;\; = v^2 - 2\tfrac{1}{n}v'\mathbf{1}_n + \tfrac{1}{n} , \qquad (60)$$
$$= \tfrac{1}{n} - v^2 .$$

Due to Lemma 7.6, all normal consensus estimates satisfy (54), thus we can apply (60) to (59) and obtain,

$$\frac{1}{n} - \hat{v}_i(t_1^{ij}(+))^2 \le \min\{\tfrac{1}{n} - \hat{v}_i(t_1^{ij})^2 , \;\; \tfrac{1}{n} - \hat{v}_j(t_0^{ij})^2\}. \qquad (61)$$

Subtracting both sides of (61) from $\frac{1}{n}$ then yields,

$$\hat{v}_i(t_1^{ij}(+))^2 \;\; \ge \max\{\hat{v}_i(t_1^{ij})^2 , \;\; \hat{v}_j(t_0^{ij})^2\}, \qquad (62)$$
$$\ge \hat{v}_i(t_1^{ij})^2 ,$$

thus each magnitude $\hat{v}_i(t)^2$ is a non-decreasing function of $t \ge 0$ for all $i \in \mathcal{V}$. ∎

Lemma 7.8: (Equality of Normalized Linear Dependent Vectors) If two linearly dependent vectors $\hat{v}_i, \hat{v}_j \in \mathbb{R}^n$ both satisfy (54), then $\hat{v}_j^2 = \hat{v}_i'\hat{v}_j = \hat{v}_i^2$.

Proof: If $\hat{v}_i$ and $\hat{v}_j$ are linearly dependent then there exists some $k \ne 0$ such that $\hat{v}_i = k\hat{v}_j$, thus if both vectors also satisfy (54) then,

$$\tfrac{1}{n}\hat{v}_i'\mathbf{1}_n = k\tfrac{1}{n}\hat{v}_j'\mathbf{1}_n \;\Rightarrow\; k = \left(\hat{v}_j'\mathbf{1}_n\right)^{-1}\left(\hat{v}_i'\mathbf{1}_n\right)$$
$$= \hat{v}_i^2/\hat{v}_j^2, \qquad (63)$$
$$\hat{v}_i^2 = k^2\hat{v}_j^2 \;\Rightarrow\; k^2 = \hat{v}_i^2/\hat{v}_j^2 .$$

Combining the RHS of the first and second lines in (63) implies $k^2 = k$ and thus $k = \pm 1$ since $k \ne 0$. We then obtain $k = 1$ since $\hat{v}_i^2/\hat{v}_j^2 > 0$, thus $\hat{v}_i = \hat{v}_j$ and the result follows. ∎

Lemma 7.9: (DA Local Zero Error Property) Every normal consensus estimate $\hat{v}_i(t)$ satisfies

$$\hat{v}_{ii}(t) = \frac{1}{n} , \;\; \forall\, i \in \mathcal{V} , \;\; \forall\, t \ge 0 . \qquad (64)$$

Proof: By Lemma 7.2, $\hat{v}_{ii}(0) = \frac{1}{n}$ for each $i \in \mathcal{V}$. Next observe that under $(A7)$ the estimate $\hat{v}_i(t)$ will not change unless a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received at node i. If a signal is received then $\hat{v}_i(t)$ is updated to the solution of (53). We now show that under the assumption (64) the solution $\hat{v}_i(t_1^{ij}(+))$ to (53) will imply (64) for every set vectors $\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$. First assume $\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij})$, and e_i are linearly dependent. In this case the update

problem (53) reduces to the RHS of (47) and thus implies (64). Next assume that any two vectors in the set $\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$ are linearly dependent. In this case (53) reduces to

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) &= \arg_{\tilde{v}} \; \min_{\tilde{v} \in \text{span}\{e_i, v\}} \; \left(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\right)^2, \\
&= \arg_{\tilde{v} = a\frac{1}{n}e_i + bv} \min_{(a,b)} \; \left(a\tfrac{1}{n}e_i + bv - \tfrac{1}{n}\mathbf{1}_n\right)^2 \;,
\end{aligned}
\tag{65}
$$

where $v \in \{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij})\}$ is linearly independent of e_i. The objective function in (65) is

$$
\begin{aligned}
f(a,b) &= \left(a\tfrac{1}{n}e_i + bv - \tfrac{1}{n}\mathbf{1}_n\right)^2 \;, \\
&= a^2(\tfrac{1}{n}e_i)^2 + b^2 v^2 + 2ab\tfrac{1}{n}e_i' v \\
&\quad - 2a\tfrac{1}{n}e_i'\tfrac{1}{n}\mathbf{1}_n - 2bv'\tfrac{1}{n}\mathbf{1}_n + \tfrac{1}{n} \;.
\end{aligned}
\tag{66}
$$

The Lemma 7.6 implies that v satisfies (54). Note also that $\frac{1}{n}e_i$ satisfies (54), thus the objective function (66) can be simplified,

$$
f(a,b) = (a^2 - 2a)\left(\frac{1}{n}e_i\right)^2 + (b^2 - 2b)v^2 + 2ab\frac{1}{n}e_i'v + \frac{1}{n} \;.
$$

The first-order partial derivatives of $f(a,b)$ are,

$$
\begin{aligned}
\frac{\partial f(a,b)}{\partial a} &= 2(a-1)\left(\tfrac{1}{n}e_i\right)^2 + 2b\tfrac{1}{n}e_i'v \;, \\
\frac{\partial f(a,b)}{\partial b} &= 2(b-1)v^2 + 2a\tfrac{1}{n}e_i'v \;.
\end{aligned}
\tag{67}
$$

The second-order partial derivatives of $f(a,b)$ are,

$$
\begin{aligned}
\frac{\partial^2 f(a,b)}{\partial a^2} &= 2\left(\tfrac{1}{n}e_i\right)^2 \;, \\
\frac{\partial^2 f(a,b)}{\partial b^2} &= 2v^2 \;, \\
\frac{\partial^2 f(a,b)}{\partial a \partial b} &= \frac{\partial^2 f(a,b)}{\partial b \partial a} = 2\tfrac{1}{n}e_i'v \;.
\end{aligned}
$$

The determinant of the Hessian matrix of $f(a,b)$ is thus,

$$
|H\big(f(a,b)\big)| = 2\left(\left(\frac{1}{n}e_i\right)^2 v^2 - \left(\frac{1}{n}e_i'v\right)^2\right) \;.
\tag{68}
$$

Since $\frac{1}{n}e_i$ and v are linearly independent, the determinant (68) is strictly positive by the Cauchy-Schwartz inequality. This implies the Hessian matrix $H\big(f(a,b)\big)$ is positive-definite, thus setting the RHS of (67) to zero and solving for (a,b) yields the unique optimal values $(\hat{a}, \hat{b})$,

$$
\hat{a} = \frac{v^2 \frac{1}{n}e_i'(\frac{1}{n}e_i - v)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2} \;, \quad \hat{b} = \frac{(\frac{1}{n}e_i)^2 v'(v - \frac{1}{n}e_i)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2} \;.
\tag{69}
$$

From (69) the unique solution to (65) is thus obtained,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) = \; &\frac{v^2 \frac{1}{n}e_i'(\frac{1}{n}e_i - v)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2}\frac{1}{n}e_i \\
&+ \frac{(\frac{1}{n}e_i)^2 v'(v - \frac{1}{n}e_i)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2}v \;.
\end{aligned}
\tag{70}
$$

Based on (70) the element $\hat{v}_{ii}(t_1^{ij}(+))$ can be expressed,

$$
\begin{aligned}
\hat{v}_{ii}(t_1^{ij}(+)) &= \frac{v^2 \frac{1}{n}e_i'(\frac{1}{n}e_i - v)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2}\frac{1}{n} \\
&\quad + \frac{(\frac{1}{n}e_i)^2 v'(v - \frac{1}{n}e_i)}{(\frac{1}{n}e_i)^2 v^2 - (\frac{1}{n}e_i'v)^2}v_i \;, \\
&= \frac{v^2(\frac{1}{n^2} - \frac{1}{n}v_i)\frac{1}{n} + \frac{1}{n^2}(v^2 - \frac{1}{n}v_i)v_i}{\frac{1}{n^2}v^2 - \frac{1}{n^2}v_i^2} \;, \\
&= \frac{\frac{1}{n}v^2 - v_i v^2 + v_i v^2 - \frac{1}{n}v_i^2}{v^2 - v_i^2} = \frac{1}{n}.
\end{aligned}
\tag{71}
$$

Note that the last equality in (71) follows since Lemma 7.8 implies $v^2 \neq v_i^2$ under the assumption that v satisfies (54) and is linearly independent from $\frac{1}{n}e_i$. Next assume that $\hat{v}_i(t_1^{ij})$, $\hat{v}_j(t_0^{ij})$ and e_i are linearly independent. In this case the solution (17) can be expressed,

$$\hat{v}_i(t_1^{ij}(+)) = V_{(DA)}(V_{(DA)}'V_{(DA)})^{-1}V_{(DA)}'\frac{1}{n}\mathbf{1}_n,$$
$$V_{(DA)} = [\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), \frac{1}{n}e_i] \ .$$

For notational convenience we denote $\hat{v}_i = \hat{v}_i(t_1^{ij})$ and $\hat{v}_j = \hat{v}_j(t_0^{ij})$. Note that $(V_{(DA)}'V_{(DA)})$ has the inverse (72) below,

$$
\begin{aligned}
(V_{(DA)}'V_{(DA)})^{-1} &= \begin{pmatrix} \hat{v}_i^2 & \hat{v}_i'\hat{v}_j & \frac{1}{n}\hat{v}_{ii} \\ \hat{v}_i'\hat{v}_j & \hat{v}_j^2 & \frac{1}{n}\hat{v}_{ji} \\ \frac{1}{n}\hat{v}_{ii} & \frac{1}{n}\hat{v}_{ji} & \frac{1}{n^2} \end{pmatrix}^{-1} , \\
&= \frac{1}{|V_{(DA)}'V_{(DA)}|} \begin{pmatrix} \frac{1}{n^2}\left(\hat{v}_j^2 - \hat{v}_{ji}^2\right) & \frac{1}{n^2}\left(\hat{v}_{ii}\hat{v}_{ji} - \hat{v}_i'\hat{v}_j\right) & \frac{1}{n}\left(\hat{v}_{ji}\hat{v}_i'\hat{v}_j - \hat{v}_j^2\hat{v}_{ii}\right) \\ \frac{1}{n^2}\left(\hat{v}_{ii}\hat{v}_{ji} - \hat{v}_i'\hat{v}_j\right) & \frac{1}{n^2}\left(\hat{v}_i^2 - \hat{v}_{ii}^2\right) & \frac{1}{n}\left(\hat{v}_{ii}\hat{v}_i'\hat{v}_j - \hat{v}_i^2\hat{v}_{ji}\right) \\ \frac{1}{n}\left(\hat{v}_{ji}\hat{v}_i'\hat{v}_j - \hat{v}_j^2\hat{v}_{ii}\right) & \frac{1}{n}\left(\hat{v}_{ii}\hat{v}_i'\hat{v}_j - \hat{v}_i^2\hat{v}_{ji}\right) & \left(\hat{v}_i^2\hat{v}_j^2 - (\hat{v}_i'\hat{v}_j)^2\right) \end{pmatrix}
\end{aligned}
\tag{72}
$$

where the determinant $|V_{(DA)}'V_{(DA)}|$ can be computed as,

$$|V_{(DA)}'V_{(DA)}| = \frac{1}{n^2}(\hat{v}_i^2\hat{v}_j^2 - (\hat{v}_i'\hat{v}_j)^2 + 2\hat{v}_{ii}\hat{v}_{ji}\hat{v}_i'\hat{v}_j$$
$$-\hat{v}_{ii}^2\hat{v}_j^2 - \hat{v}_i^2\hat{v}_{ji}^2) \ .$$

Right-multiplying (72) by $V_{(DA)}'\frac{1}{n}\mathbf{1}_n$ and then left-multiplying by the first row of $V_{(DA)}$ yields $\hat{v}_{ii}(t_1^{ij}(+))$,

$$
\begin{aligned}
\hat{v}_{ii}(t_1^{ij}(+)) &= \frac{1}{|V_{(DA)}'V_{(DA)}|}\frac{1}{n^2}\big(\hat{v}_i^2\hat{v}_j^2\hat{v}_{ii} - \hat{v}_i^2\hat{v}_{ji}^2\hat{v}_{ii} \\
&\quad +\hat{v}_{ii}^2\hat{v}_{ji}\hat{v}_j^2 - \hat{v}_{ii}\hat{v}_i'\hat{v}_j\hat{v}_j^2 + \frac{1}{n}\hat{v}_{ji}\hat{v}_i'\hat{v}_j\hat{v}_{ii} - \frac{1}{n}\hat{v}_{ii}^2\hat{v}_j^2 \\
&\quad +\hat{v}_i^2\hat{v}_{ii}\hat{v}_{ji}^2 - \hat{v}_i'\hat{v}_j\hat{v}_i^2\hat{v}_{ji} + \hat{v}_i^2\hat{v}_j^2\hat{v}_{ji} - \hat{v}_{ii}^2\hat{v}_j^2\hat{v}_{ji} \\
&\quad +\frac{1}{n}\hat{v}_{ii}\hat{v}_{ji}\hat{v}_i'\hat{v}_j - \frac{1}{n}\hat{v}_i^2\hat{v}_{ji}^2 + \hat{v}_i^2\hat{v}_i'\hat{v}_j\hat{v}_{ji} - \hat{v}_i^2\hat{v}_j^2\hat{v}_{ii} \\
&\quad +\hat{v}_{ii}\hat{v}_i'\hat{v}_j\hat{v}_j^2 - \hat{v}_i^2\hat{v}_j^2\hat{v}_{ji} + \frac{1}{n}\hat{v}_i^2\hat{v}_j^2 - \frac{1}{n}(\hat{v}_i'\hat{v}_j)^2\big) \ , \\
&= \frac{1}{|V_{(DA)}'V_{(DA)}|}\frac{1}{n^3}\big(\hat{v}_i^2\hat{v}_j^2 - (\hat{v}_i'\hat{v}_j)^2 \\
&\quad +2\hat{v}_{ii}\hat{v}_{ji}\hat{v}_i'\hat{v}_j - \hat{v}_{ii}^2\hat{v}_j^2 - \hat{v}_i^2\hat{v}_{ji}^2\big) \ , \\
&= \frac{1}{n} \ .
\end{aligned}
\tag{73}
$$

The last line in (73) follows since $|V_{(DA)}'V_{(DA)}|$ is non-zero if $e_i, \hat{v}_i$ and $\hat{v}_j$ are linearly independent. ∎

Lemma 7.10: (DA Lower Bound On Signal Reduction in Error) Upon reception of any signal $S^{ij}(t_0^{ij}, t_1^{ij})$ the decrease in the updated normal consensus squared error has the following lower bound,

$$
\begin{aligned}
E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij})\big) - E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij}(+))\big) &\geq \\
\max\{&\hat{v}_j(t_0^{ij})^2 - \hat{v}_i(t_1^{ij})^2 \ , \\
&n\big(\hat{v}_j(t_0^{ij})'(\hat{v}_j(t_0^{ij}) - \hat{v}_i(t_1^{ij}))\big)^2\}.
\end{aligned}
\tag{74}
$$

Proof: By Lemma 7.6 we can apply (60) to the left-hand side (LHS) of (74),

$$E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij})\big) - E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij}(+))\big) = \hat{v}_i(t_1^{ij}(+))^2 - \hat{v}_i(t_1^{ij})^2. \tag{75}$$

Applying the first line of (62) to the RHS of (75) then yields,

$$
\begin{aligned}
E^2_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t_1^{ij})\big) &- E^2_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t_1^{ij}(+))\big) \\
&\geq \max\{\hat{v}_i(t_1^{ij})^2, \hat{v}_j(t_0^{ij})^2\} - \hat{v}_i(t_1^{ij})^2 \ , \\
&\geq \max\{0 \ , \ \hat{v}_j(t_0^{ij})^2 - \hat{v}_i(t_1^{ij})^2\} \ , \\
&\geq \hat{v}_j(t_0^{ij})^2 - \hat{v}_i(t_1^{ij})^2 \ .
\end{aligned}
\tag{76}
$$

Next observe that for any two vectors $\hat{v}_i, \hat{v}_j$,

$$
\mathrm{span}\ \{\hat{v}_i, \hat{v}_j\} \subseteq \mathrm{span}\ \{\hat{v}_i, \hat{v}_j, e_i\} \ .
$$

It thus follows that,

$$
\begin{aligned}
E_{\frac{1}{n}\mathbf{1}_n}&\big(\mathrm{arg}_{\tilde{v}}\ \min_{\tilde{v}\, \in\, \mathrm{span}\{\hat{v}_i, \hat{v}_j, e_i\}}\ (\tilde{v} - \tfrac{1}{n}\mathbf{1}_n)^2 \ \big) \\
&\leq E_{\frac{1}{n}\mathbf{1}_n}\big(\mathrm{arg}_{\tilde{v}}\ \min_{\tilde{v}\, \in\, \mathrm{span}\{\hat{v}_i, \hat{v}_j\}}\ (\tilde{v} - \tfrac{1}{n}\mathbf{1}_n)^2 \ \big).
\end{aligned}
\tag{77}
$$

Let $\hat{v}_i = \hat{v}_i(t_1^{ij})$, $\hat{v}_j = \hat{v}_j(t_0^{ij})$ and $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(+)$ for notational convenience. From (77) we have,

$$
\begin{aligned}
E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i) &- E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(+)) \\
&\geq \big(\mathrm{arg}_{\tilde{v}}\ \min_{\tilde{v}\, \in\, \mathrm{span}\{\hat{v}_i, \hat{v}_j\}}\ (\tilde{v} - \tfrac{1}{n}\mathbf{1}_n)^2\big)^2 - \hat{v}_i^2 \ , \\
&= \hat{w}^2 - \hat{v}_i^2 \ , \\
&= \tfrac{1}{n}\big(\hat{w}'\mathbf{1}_n - \hat{v}_i'\mathbf{1}_n\big) \ ,
\end{aligned}
\tag{78}
$$

where the last line holds due to Lemma 7.6, and $\hat{w}$ is defined,

$$
\begin{aligned}
\hat{w}\ &= \mathrm{arg}_{\tilde{v}}\ \min_{\tilde{v}\, \in\, \mathrm{span}\{\hat{v}_i, \hat{v}_j\}}\ (\tilde{v} - \tfrac{1}{n}\mathbf{1}_n)^2 \ , \\
&= \mathrm{arg}_{\tilde{v}\, =\, a\hat{v}_i + b\hat{v}_j}\ \min_{(a, b)}\ (a\hat{v}_i + b\hat{v}_j - \tfrac{1}{n}\mathbf{1}_n)^2 \ .
\end{aligned}
\tag{79}
$$

Note that Lemma 7.7 together with the initialization (45) implies $\hat{v}_i$ and $\hat{v}_j$ are non-zero. Next assume that $\hat{v}_i$ and $\hat{v}_j$ are linearly dependent. In this case (79) reduces to,

$$
\begin{aligned}
\hat{w}\ &= \mathrm{arg}_{\tilde{v}}\ \min_{\tilde{v}\, \in\, \mathrm{span}\{\hat{v}_i\}}\ (\tilde{v} - \tfrac{1}{n}\mathbf{1}_n)^2 \ , \\
&= V(V'V)^{-1}V'\tfrac{1}{n}\mathbf{1}_n \ , \ V = [\hat{v}_i] \ , \\
&= \hat{v}_i \tfrac{\hat{v}_i'\mathbf{1}_n}{n\hat{v}_i^2} \ , \\
&= \hat{v}_i \ ,
\end{aligned}
\tag{80}
$$

where the last line follows due to Lemma 7.6. Applying (80) to (78) implies,

$$
\begin{aligned}
E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i) - E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(+))\ &\geq \hat{w}^2 - \hat{v}_i^2 \ , \\
&= \hat{v}_i^2 - \hat{v}_i^2 \\
&= 0 \ , \\
&= n\big(\hat{v}_j^2 - \hat{v}_i'\hat{v}_j\big)^2
\end{aligned}
\tag{81}
$$

where the last line follows by Lemma 7.8 since Lemma 7.6 implies both $\hat{v}_i$ and $\hat{v}_j$ satisfy (54), and we are assuming $\hat{v}_i$ and $\hat{v}_j$ are linearly dependent.

Next assume $\hat{v}_i$ and $\hat{v}_j$ are linearly independent. In this case (79) can be solved analogous to (70) based on the optimization problem (65),

$$\begin{aligned}
\hat{w} \quad &= \arg_{\tilde{v} \, = \, a\hat{v}_i + b\hat{v}_j} \ \min_{(a,b)} \ \left(a\hat{v}_i + b\hat{v}_j - \tfrac{1}{n}\mathbf{1}_n\right)^2 \ , \\
&= \frac{\hat{v}_j^2 \hat{v}_i'(\hat{v}_i - \hat{v}_j)}{\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2}\hat{v}_i + \frac{\hat{v}_i^2 \hat{v}_j'(\hat{v}_j - \hat{v}_i)}{\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2}\hat{v}_j \ .
\end{aligned} \tag{82}$$

Substituting the second line of (82) for $\hat{w}$ in the third line of (78) then yields,

$$\begin{aligned}
\tfrac{1}{n}&\left(\hat{w}'\mathbf{1}_n - \hat{v}_i'\mathbf{1}_n\right) \\
&= \tfrac{1}{n}\left(\frac{\hat{v}_j^2 \hat{v}_i'(\hat{v}_i - \hat{v}_j)\hat{v}_i'\mathbf{1}_n + \hat{v}_i^2 \hat{v}_j'(\hat{v}_j - \hat{v}_i)\hat{v}_j'\mathbf{1}_n}{\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2} - \hat{v}_i'\mathbf{1}_n\right) \\
&= \frac{\hat{v}_j^2 \hat{v}_i^2 \hat{v}_i'(\hat{v}_i - \hat{v}_j) + \hat{v}_i^2 \hat{v}_j^2 \hat{v}_j'(\hat{v}_j - \hat{v}_i) - \left(\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2\right)\hat{v}_i^2}{\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2} \\
&= \frac{\hat{v}_i^2}{\hat{v}_i^2 \hat{v}_j^2 - (\hat{v}_i' \hat{v}_j)^2}\left(\hat{v}_j^2(\hat{v}_i^2 - \hat{v}_i'\hat{v}_j) + \hat{v}_j^2(\hat{v}_j^2 - \hat{v}_i'\hat{v}_j)\right. \\
&\hspace{7cm}\left. - \hat{v}_i^2\hat{v}_j^2 + (\hat{v}_i'\hat{v}_j)^2\right), \\
&\geq \left(\hat{v}_j^4 - 2\hat{v}_i'\hat{v}_j\hat{v}_j^2 + (\hat{v}_i'\hat{v}_j)^2\right)\frac{\hat{v}_i^2}{\hat{v}_i^2\hat{v}_j^2} \\
&= \left(\hat{v}_j^2 - \hat{v}_i'\hat{v}_j\right)^2\frac{1}{\hat{v}_j^2} \\
&\geq n\left(\hat{v}_j^2 - \hat{v}_i'\hat{v}_j\right)^2 \ ,
\end{aligned} \tag{83}$$

where the last line follows since $E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_j) \geq 0$ implies $\hat{v}_j^2 \leq \frac{1}{n}$. Combining (78) and (83) implies,

$$E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i) - E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(+)) \geq n\left(\hat{v}_j^2 - \hat{v}_i'\hat{v}_j\right)^2 \ . \tag{84}$$

Together (76), (81) and (84) imply (74). ∎

Lemma 7.11: (DA Non-Decreasing Magnitude for any Communication Path) Any communication path $C^{ij}_{[t_0(ij), t_1(ij)]} \in C_{[0, t_1]}$ implies,

$$\hat{v}_i(t_1(ij)(+))^2 \geq \hat{v}_j(t_0(ij))^2 \ . \tag{85}$$

Proof: The first line in (62) implies that for any signal $S^{ij}(t_0^{ij}, t_1^{ij})$,

$$\hat{v}_i(t_1^{ij}(+))^2 \geq \hat{v}_j(t_0^{ij})^2 \ .$$

Thus for any communication path $C^{ij}_{[t_0(ij), t_1(ij)]}$ defined as in (33),

$$\begin{aligned}
\hat{v}_i(t_1^{i\ell_{k(ij)}}(+))^2 &\geq \hat{v}_{\ell_{k(ij)}}(t_0^{i\ell_{k(ij)}})^2 \\
&\geq \hat{v}_{\ell_{k(ij)}}(t_1^{\ell_{k(ij)}\ell_{k(ij)-1}}(+))^2 \\
&\geq \hat{v}_{\ell_{k(ij)-1}}(t_0^{\ell_{k(ij)}\ell_{k(ij)-1}})^2 \\
&\geq \hat{v}_{\ell_{k(ij)-1}}(t_1^{\ell_{k(ij)-1}\ell_{k(ij)-2}}(+))^2 \geq \cdots \\
&\geq \hat{v}_{\ell_1}(t_0^{\ell_2\ell_1})^2 \geq \ \hat{v}_{\ell_1}(t_1^{\ell_1 j}(+))^2 \geq \hat{v}_j(t_0^{\ell_1 j})^2 \ ,
\end{aligned}$$

where every second inequality holds due to Lemma 7.7 since $t_0^{\ell_{q+1}\ell_q} > t_1^{\ell_q\ell_{q-1}}$ for each $q = 1, 2, \ldots, k(ij)$, where $\ell_0 = j$ and $\ell_{k(ij)+1} = i$. We then obtain (85) again by Lemma 7.7 since $\hat{v}_i(t_1(ij)(+))^2 \geq \hat{v}_i(t_1^{i\ell_{k(ij)}}(+))^2$ and $\hat{v}_j(t_0^{\ell_1 j})^2 \geq \hat{v}_j(t_0(ij))^2$ because $t_1(ij) \geq t_1^{i\ell_{k(ij)}}$ and $t_0(ij) \leq t_0^{\ell_1 j}$ respectively. ∎

Lemma 7.12: (Error Expression for $C_{[0,t_1]}$ satisfying (35)) For any communication sequence $C_{[0,t_1]}$ satisfying (35) the total reduction in normal consensus squared error $\mathbf{E}^2(C_{[0,t_1]})$ defined in (42) is,

$$
\begin{aligned}
\mathbf{E}^2(C_{[0,t_1]}) &= \sum_{i=1}^n \left(E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(0)) - E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(t_1(+))) \right) \\
&= \frac{n-1}{n} - \sum_{i=1}^n E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(t_1(+))) \\
&= \sum_{\ell \in \mathbb{N}} \mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]}) \ , \ C_{[t_0^\ell, t_1^\ell]} \in \mathcal{SVSC} \\
&\leq \frac{n-1}{n} \ .
\end{aligned}
\tag{86}
$$

Proof: The first line follows from (A7),(A8) and the definitions $(41) - (43)$. The second line in (86) is due to the initialization (45). The third line in (86) follows since any communication sequence $C_{[0,t_1]}$ satisfying (35) can be partitioned into an infinite number of disjoint sequences $C_{[t_0^\ell, t_1^\ell]} \in \mathcal{SVSC}$. The last line in (86) follows since the minimum error of any normal consensus estimate is 0. ∎

Lemma 7.13: (Vanishing Reduction in Error for $C_{[0,t_1]}$ satisfying (35)) For any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_ε such that,

$$
\mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]}) \leq \varepsilon \ , \ \forall \ \ell \geq \ell_\varepsilon \ ,
\tag{87}
$$

for any $\varepsilon > 0$, where $\mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]})$ is defined by (42).

Proof: The third line of (86) implies that for any $C_{[0,t_1]}$ satisfying (35) the quantity $\mathbf{E}^2(C_{[0,t_1]})$ is the sum of an infinite number of non-negative terms $\mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]})$, the fourth line in (86) implies $\mathbf{E}^2(C_{[0,t_1]})$ is bounded above, thus (87) follows by the monotonic sequence theorem. ∎

Lemma 7.14: (DA Lower Bound on Reduction in Error for any $C_{[0,t_1]}$ satisfying (34)) Any communication sequence $C_{[t_0,t_1]}$ satisfying (34) implies,

$$
\begin{aligned}
\mathbf{E}^2(C_{[t_0,t_1]}) &\geq n \big(\min_{i \in \mathcal{V}} \{\hat{v}_i(t_1(+))^2\} \\
&\qquad - \max_{i \in \mathcal{V}} \{\hat{v}_i(t_0)^2\} \big) \ , \\
&\geq 0 \ .
\end{aligned}
\tag{88}
$$

Proof: The first line in (88) holds for any communication sequence $C_{[t_0,t_1]}$,

$$
\begin{aligned}
\mathbf{E}^2(C_{[t_0,t_1]}) &= \sum_{i=1}^n \left(E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(t_0)) - E^2_{\frac{1}{n}\mathbf{1}_n}(\hat{v}_i(t_1(+))) \right) \\
&= \sum_{i=1}^n \left(\hat{v}_i(t_1(+))^2 - \hat{v}_i(t_0)^2 \right) \\
&\geq n \big(\min_{i \in \mathcal{V}} \{\hat{v}_i(t_1(+))^2\} - \max_{i \in \mathcal{V}} \{\hat{v}_i(t_0)^2\} \big) \ .
\end{aligned}
$$

To prove the second line in (88) it is required that $C_{[t_0,t_1]}$ satisfies (34). We define,

$$
\begin{aligned}
\underline{\ell} &= \arg_i \ \min_{i \in \mathcal{V}} \{\hat{v}_i(t_1(+))^2\} \ , \\
\overline{\ell} &= \arg_i \ \max_{i \in \mathcal{V}} \{\hat{v}_i(t_0)^2\} \ .
\end{aligned}
$$

Since $C_{[t_0,t_1]}$ satisfies (34) there exists a communication path $C^{\underline{\ell}\overline{\ell}}_{[t_0(\underline{\ell}\overline{\ell}), t_1(\underline{\ell}\overline{\ell})]} \in C_{[t_0,t_1]}$. The Lemma 7.11 then implies,

$$
\hat{v}_{\underline{\ell}}(t_1(\underline{\ell}\overline{\ell})(+))^2 \geq \hat{v}_{\overline{\ell}}(t_0(\underline{\ell}\overline{\ell}))^2 \ .
$$

The second line in (88) then follows because,

$$
\begin{aligned}
\min_{i \in \mathcal{V}}\{\hat{v}_i(t_1(+))^2\} &= \hat{v}_{\underline{\ell}}(t_1(+))^2 \\
&\geq \hat{v}_{\underline{\ell}}(t_1(\underline{\ell}\overline{\ell})(+))^2 \\
&\geq \hat{v}_{\overline{\ell}}(t_0(\underline{\ell}\overline{\ell}))^2 \\
&\geq \hat{v}_{\overline{\ell}}(t_0)^2 = \max_{i \in \mathcal{V}}\{\hat{v}_i(t_0)^2\} \ ,
\end{aligned}
$$

where the first and third inequality hold due to Lemma 7.7 because $t_1 \geq t_1(\underline{\ell}\overline{\ell})$ and $t_0(\underline{\ell}\overline{\ell}) \geq t_0$ respectively. ∎

Lemma 7.15: (DA Local Convergence of Normal Consensus Estimates) For any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_χ such that for all $\ell \geq \ell_\chi$,

$$
\left(\hat{v}_i(t_1^{ij}) - \hat{v}_j(t_0^{ij})\right)^2 \leq \chi \ , \ \forall \ S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]} \ , \tag{89}
$$

for any $\chi > 0$.

Proof: For any communication sequence $C_{[0,t_1]}$ satisfying (35) the Lemma 7.13 implies there exists an integer ℓ_ε such that (87) holds for any $\varepsilon > 0$. For all $\ell \geq \ell_\varepsilon$ we thus have for any signal $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij})^2 - \hat{v}_j(t_0^{ij})^2 &\leq \hat{v}_i(t_1^\ell(+))^2 - \hat{v}_j(t_0^\ell)^2 \ , \\
&\leq \sum_{r=1}^n \left(\hat{v}_r(t_1^\ell(+))^2 - \hat{v}_r(t_0^\ell)^2\right) \ , \\
&= \mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]}) \leq \varepsilon \ .
\end{aligned} \tag{90}
$$

Note that the first inequality in (90) holds by Lemma 7.7 since $t_0^{ij} \geq t_0^\ell$ and $t_1^{ij} \leq t_1^\ell$ for any $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$. The second inequality in (90) holds since,

$$
\begin{aligned}
&\sum_{r=1}^n \left(\hat{v}_r(t_1^\ell(+))^2 - \hat{v}_r(t_0^\ell)^2\right) \\
&= \left(\hat{v}_i(t_1^\ell(+))^2 - \hat{v}_j(t_0^\ell)^2\right) + \left(\hat{v}_j(t_1^\ell(+))^2 - \hat{v}_i(t_0^\ell)^2\right) \\
&\quad + \sum_{r \in \mathcal{V}\backslash\{i,j\}} \left(\hat{v}_r(t_1^\ell(+))^2 - \hat{v}_r(t_0^\ell)^2\right) \ ,
\end{aligned} \tag{91}
$$

where

$$
\sum_{r \in \mathcal{V}\backslash\{i,j\}} \left(\hat{v}_r(t_1^\ell(+))^2 - \hat{v}_r(t_0^\ell)^2\right) \geq 0 \tag{92}
$$

holds due to Lemma 7.7, and,

$$
\begin{aligned}
\hat{v}_j(t_1^\ell(+))^2 - \hat{v}_i(t_0^\ell)^2 &\geq \min_{r \in \mathcal{V}}\{\hat{v}_r(t_1^\ell(+))^2\} \\
&\quad - \max_{r \in \mathcal{V}}\{\hat{v}_r(t_0^\ell)^2\} \\
&\geq 0 \ ,
\end{aligned} \tag{93}
$$

where the second inequality in (93) holds due to Lemma 7.14 since $C_{[t_0^\ell, t_1^\ell]} \in \mathcal{SVSC}$ and thus satisfies (34). Together (91), (92) and (93) imply the second inequality in (90). Applying Lemma 7.10 to Lemma 7.13 implies that for all $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$ and $\ell \geq \ell_\varepsilon$,

$$
\begin{aligned}
\varepsilon &\geq \mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]}) \\
&\geq \mathbf{E}^2\left(S^{ij}(t_0^{ij}, t_1^{ij})\right) \ , \\
&= E^2_{\frac{1}{n}\mathbf{1}_n}\left(\hat{v}_i(t_1^{ij})\right) - E^2_{\frac{1}{n}\mathbf{1}_n}\left(\hat{v}_i(t_1^{ij}(+))\right) \ , \\
&\geq \max\{\hat{v}_j(t_0^{ij})^2 - \hat{v}_i(t_1^{ij})^2 \ , \\
&\qquad n\left(\hat{v}_j(t_0^{ij})'(\hat{v}_j(t_0^{ij}) - \hat{v}_i(t_1^{ij}))\right)^2\} \ ,
\end{aligned} \tag{94}
$$

for any $\varepsilon > 0$. For notational convenience denote $\hat{v}_i = \hat{v}_i(t_1^{ij})$ and $\hat{v}_j = \hat{v}_j(t_0^{ij})$. Combining (90) and (94) implies that for any $\varepsilon > 0$ there exist an integer ℓ_ε such that,

$$\begin{aligned}
\hat{v}_i^2 - \hat{v}_j^2 &\leq \varepsilon \; , \\
\hat{v}_j'(\hat{v}_j - \hat{v}_i) &\leq \sqrt{\varepsilon/n} \; ,
\end{aligned} \tag{95}$$

for any $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$ and $\ell \geq \ell_\varepsilon$. To obtain (89) note that (95) implies,

$$\begin{aligned}
(\hat{v}_i - \hat{v}_j)^2 &= \hat{v}_i^2 - 2\hat{v}_i'\hat{v}_j + \hat{v}_j^2 \; , \\
&= \hat{v}_i^2 - \hat{v}_j^2 + 2\hat{v}_j'(\hat{v}_j - \hat{v}_i) \; , \\
&\leq \varepsilon + 2\sqrt{\varepsilon/n} \; , \\
&\leq \sqrt{\varepsilon}(1 + 2/\sqrt{n}) \; , \; \forall \; \epsilon \in (0, 1].
\end{aligned} \tag{96}$$

We thus define $\overline{\varepsilon}(\chi)$,

$$\overline{\varepsilon}(\chi) = \left(\frac{\chi}{1 + 2/\sqrt{n}} \right)^2 \; . \tag{97}$$

For any $\chi \in (0, 1)$ and $\varepsilon \in (0, \overline{\varepsilon}(\chi)]$ the result (89) then follows from (96). ∎

Lemma 7.16: (DA Properties of the Normal Consensus Update) Upon reception of any signal $S^{ij}(t_0^{ij}, t_1^{ij})$ the normal consensus estimate $\hat{v}_i(t_1^{ij}(+))$ that results from the update problem (53) will satisfy,

$$\hat{v}_i(t_1^{ij})'\big(\hat{v}_i(t_1^{ij}) - \hat{v}_i(t_1^{ij}(+))\big) \leq 0 \; . \tag{98}$$

Proof: Let us define $\tilde{w}$,

$$\tilde{w} = \arg_{\tilde{v}} \min_{\tilde{v} \in \text{span}\{\hat{v}_i(t_1^{ij}(+)), \hat{v}_i(t_1^{ij}), e_i\}} \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n \right)^2 \; , \tag{99}$$

where $\hat{v}_i(t_1^{ij}(+))$ is given by (53). Note that $\hat{v}_i(t_1^{ij}(+)) \in \text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$ implies,

$$\text{span}\{\hat{v}_i(t_1^{ij}(+)), \hat{v}_i(t_1^{ij}), e_i\} \subseteq \text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\} \; . \tag{100}$$

From (100) we have,

$$E_{\frac{1}{n}\mathbf{1}_n}(\tilde{w}) \geq E_{\frac{1}{n}\mathbf{1}_n}\big(\hat{v}_i(t_1^{ij}(+))\big) \; , \tag{101}$$

and thus combining (101) and (60) implies,

$$\tilde{w}^2 \leq \hat{v}_i(t_1^{ij}(+))^2 \; . \tag{102}$$

Next observe that since $\tilde{w}$ is defined by (99), if (102) is applied to the signal $S^{ii}(t_1^{ij}, t_1^{ij}(+))$ then,

$$\begin{aligned}
\mathbf{E}^2\big(S^{ii}(t_1^{ij}, t_1^{ij}(+))\big) &\equiv E_{\frac{1}{n}\mathbf{1}_n}^2\big(\hat{v}_i(t_1^{ij}(+))\big) - E_{\frac{1}{n}\mathbf{1}_n}^2(\tilde{w}) \; , \\
&= \tilde{w}^2 - \hat{v}_i(t_1^{ij}(+))^2 \leq 0 \; .
\end{aligned} \tag{103}$$

Applying Lemma 7.10 to the signal $S^{ii}(t_1^{ij}, t_1^{ij}(+))$ then implies,

$$\begin{aligned}
0 &\geq \mathbf{E}^2\big(S^{ii}(t_1^{ij}, t_1^{ij}(+))\big) \\
&\geq \max\{\hat{v}_i(t_1^{ij})^2 - \hat{v}_i(t_1^{ij}(+))^2 \; , \\
&\qquad n\big(\hat{v}_i(t_1^{ij})'(\hat{v}_i(t_1^{ij}) - \hat{v}_i(t_1^{ij}(+)))\big)^2\}
\end{aligned} \tag{104}$$

where the first line follows from (103) and the last line implies (98). ∎

Lemma 7.17: (DA Vanishing Change in Normal Consensus Update) For any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_ε such that for all $\ell \geq \ell_\varepsilon$,

$$\big(\hat{v}_i(t_1^{ij}(+)) - \hat{v}_i(t_1^{ij})\big)^2 \leq \varepsilon \ , \ \forall \ S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]} \ , \tag{105}$$

for any $\varepsilon > 0$.

Proof: Recall that Lemma 7.13 implies that for any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_ε such that (87) holds for any $\varepsilon > 0$, we thus observe for $\ell \geq \ell_\varepsilon$,

$$\begin{aligned} \varepsilon \ &\geq \mathbf{E}^2\big(C_{[t_0^\ell, t_1^\ell]}\big) \\ &\geq \mathbf{E}^2(S^{ij}(t_0^{ij}, t_1^{ij})) \ , \ \forall \ S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]} \\ &= \hat{v}_i(t_1^{ij}(+))^2 - \hat{v}_i(t_1^{ij})^2 \ . \end{aligned} \tag{106}$$

For all $\ell \geq \ell_\varepsilon$ and signals $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$ we then have,

$$\begin{aligned} \big(\hat{v}_i(t_1^{ij}(+)) - \hat{v}_i(t_1^{ij})\big)^2 &= \hat{v}_i(t_1^{ij}(+))^2 - \hat{v}_i(t_1^{ij})^2 \\ &\quad + 2\hat{v}_i(t_1^{ij})'\big(\hat{v}_i(t_1^{ij}) - \hat{v}_i(t_1^{ij}(+))\big) \\ &\leq \mathbf{E}^2\big(S^{ij}(t_0^{ij}, t_1^{ij})\big) \leq \varepsilon \ , \end{aligned}$$

where the first inequality follows from Lemma 7.16 and the second inequality from Lemma 7.17. ∎

Lemma 7.18: (DA Vanishing Change between Normal Consensus Update and Signal) For any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_γ such that,

$$\big(\hat{v}_i(t_1^{ij}(+)) - \hat{v}_j(t_0^{ij})\big)^2 \leq \gamma \ , \ \forall \ S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}, \tag{107}$$

for all $\ell \geq \ell_\gamma$ and any $\gamma > 0$.

Proof: For any communication sequence $C_{[0,t_1]}$ satisfying (35) the Lemma 7.15 implies that (89) holds for any $\chi \in (0, 1)$ and $\varepsilon \in (0, \overline{\varepsilon}(\chi)]$, where $\overline{\varepsilon}(\chi)$ is given by (97). The Lemma 7.17 implies that (105) holds for any $\varepsilon > 0$, thus for any $\ell \geq \ell_{\overline{\varepsilon}(\chi)}$ and $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0^\ell, t_1^\ell]}$ the triangle inequality then implies,

$$\begin{aligned} \sqrt{\big(\hat{v}_i(t_1^{ij}(+)) - \hat{v}_j(t_0^{ij})\big)^2} &\leq \sqrt{\big(\hat{v}_i(t_1^{ij}(+)) - \hat{v}_i(t_1^{ij})\big)^2} \\ &\quad + \sqrt{\big(\hat{v}_i(t_1^{ij}) - \hat{v}_j(t_0^{ij})\big)^2} \\ &\leq \sqrt{\overline{\varepsilon}(\chi)} + \sqrt{\chi} \ , \\ &\leq 2\sqrt{\chi} \ , \ \forall \ \chi \in (0, 1). \end{aligned}$$

Any $\chi \in (0, \gamma/4]$ and $\varepsilon \in (0, \overline{\varepsilon}(\chi)]$ thus yields (107) for any $\gamma \in (0, 4]$. ∎

Lemma 7.19: (DA Network Convergence to Average-Consensus) For any communication sequence $C_{[0,t_1]}$ satisfying (35) there exists an integer ℓ_ξ such that for all $\ell \geq \ell_\xi$,

$$\sum_{i=1}^n \big(\hat{v}_i(t_1^\ell(+)) - \frac{1}{n}\mathbf{1}_n\big)^2 \leq \xi \ , \ \forall \ C_{[t_0^\ell, t_1^\ell]} \in C_{[0,t_1]}, \tag{108}$$

for any $\xi > 0$.

Proof: Since $C_{[0,t_1]}$ satisfies (35) we have $C_{[t_0^\ell,t_1^\ell]} \in \mathcal{SVSC}$ for each $\ell \in \mathbb{N}$, thus there exists a communication path $C_{[t_0(ij),t_1(ij)]}^{ij} \in C_{[t_0^\ell,t_1^\ell]}$ for any $i \in \mathcal{V}$, $j \in \mathcal{V}_{-i}$, and $\ell \in \mathbb{N}$. For any $i \in \mathcal{V}$, $j \in \mathcal{V}_{-i}$ and $\ell \in \mathbb{N}$ the triangle inequality then implies,

$$
\begin{aligned}
\sqrt{\left(\hat{v}_{ij}(t_1^\ell(+)) - \hat{v}_{jj}(t_0^{\ell_1 j})\right)^2} \\
\leq \sum_{S^{rp}(t_0^{rp},t_1^{rp}) \in C_{[t_0(ij),t_1(ij)]}^{ij}} \\
\sqrt{\left(\hat{v}_{rj}(t_1^{rp}(+)) - \hat{v}_{pj}(t_0^{rp})\right)^2} \\
+ \sum_{q=1}^{k(ij)+1} \sum_{S^{rp}(t_0^{rp},t_1^{rp}) \in Q_q^\ell(ij)} \\
\sqrt{\left(\hat{v}_{rj}(t_1^{rp}(+)) - \hat{v}_{rj}(t_1^{rp})\right)^2}
\end{aligned}
\tag{109}
$$

where we define $\bar{C}_{[t_0^\ell,t_1^\ell]} = C_{[t_0^\ell,t_1^\ell]} \setminus C_{[t_0(ij),t_1(ij)]}^{ij}$ and,

$$
\begin{aligned}
Q_q^\ell(ij) = \{S^{\ell_q m}(t_0^{\ell_q m}, t_1^{\ell_q m}) \in \bar{C}_{[t_0^\ell,t_1^\ell]} \ : \\
t_1^{\ell_q m} \in (t_1^{\ell_q \ell_{q-1}}, t_0^{\ell_{q+1} \ell_q})\}, \\
\text{for } q = 1,\ldots,k(ij), \text{ where } \ell_0 = j,\, \ell_{k(ij)+1} = i, \\
Q_{k(ij)+1}^\ell(ij) = \{S^{im}(t_0^{im}, t_1^{im}) \in C_{[t_0^\ell,t_1^\ell]} \ : \\
t_1^{im} > t_1(ij)\} \ .
\end{aligned}
\tag{110}
$$

We clarify that the RHS of (109) includes the differences between the received normal consensus vector $\hat{v}_{\ell_{q-1}}(t_0^{\ell_q \ell_{q-1}})$ and the updated normal consensus vector $\hat{v}_{\ell_q}(t_1^{\ell_q \ell_{q-1}}(+))$ that results from each signal contained in the communication path $C_{[t_0(ij),t_1(ij)]}^{ij} \in C_{[t_0^\ell,t_1^\ell]}$. Each set $Q_q^\ell(ij)$ defined in (110) contains the signals received at each node *after* the respective signal in communication path was received, but *before* the respective signal in the communication path was sent, as is required for an application of the triangle inequality. The set $Q_{k(ij)+1}^\ell(ij)$ contains the signals received at node i *after* the last signal in the communication path $C_{[t_0(ij),t_1(ij)]}^{ij}$ was received, but *before or at* the end of the communication sequence $C_{[t_0^\ell,t_1^\ell]}$, again this is required for application of the triangle inequality.

For any communication sequence $C_{[0,t_1]}$ satisfying (35) the Lemma 7.17 implies there exists an integer ℓ_ε such that (105) holds for any $\varepsilon > 0$. Likewise, for any communication sequence $C_{[0,t_1]}$ satisfying (35) the Lemma 7.18 implies there exists an integer ℓ_γ such that (107) holds for any $\gamma > 0$. Thus for any $\gamma \in (0,4]$ if we let $\chi \in (0,\gamma/4]$ and $\varepsilon \in (0,\overline{\varepsilon}(\chi)]$ then for any $\ell \geq \ell_{\overline{\varepsilon}(\chi)}$,

$$
\begin{aligned}
\sqrt{\left(\hat{v}_{ij}(t_1^\ell(+)) - \hat{v}_{jj}(t_0^{\ell_1 j})\right)^2} \\
\leq |C_{[t_0(ij),t_1(ij)]}^{ij}| \sqrt{\gamma} \\
+ \sum_{q=1}^{k(ij)+1} |Q_q^\ell(ij)| \sqrt{\overline{\varepsilon}(\chi)} \ , \\
\leq \left(|C_{[t_0(ij),t_1(ij)]}^{ij}| + \sum_{q=1}^{k(ij)+1} |Q_q^\ell(ij)|\right) \sqrt{\gamma} \ , \\
\leq |C_{[t_0^\ell,t_1^\ell]}| \sqrt{\gamma} \ ,
\end{aligned}
\tag{111}
$$

where the second inequality holds since $\overline{\varepsilon}(\chi) < \gamma$, and the last inequality holds since every signal contained in $C_{[t_0^\ell,t_1^\ell]}$ is represented by at most one term on the RHS of (109), that is,

$$
|C_{[t_0^\ell,t_1^\ell]}| \geq |C_{[t_0(ij),t_1(ij)]}^{ij}| + \sum_{q=1}^{k(ij)+1} |Q_q^\ell(ij)| \ .
$$

Due to (111), any $\xi > 0$ and $\gamma \in (0, \xi/(n|C_{[t_0^\ell, t_1^\ell]}|)^2]$ then implies,

$$
\begin{aligned}
\big(\hat{v}_{ij}(t_1^\ell(+)) - \hat{v}_{jj}(t_0^{\ell_1 j})\big)^2 \quad &= \big(\hat{v}_{ij}(t_1^\ell(+)) - \tfrac{1}{n}\big)^2 , \\
&\leq \xi/n^2 ,
\end{aligned}
\tag{112}
$$

for any $(i,j) \in \mathcal{V}^2$, where the first equality holds due to Lemma 7.9. The result (108) then follows from (112) since

$$
\begin{aligned}
\sum_{i=1}^n \big(\hat{v}_i(t_1^\ell(+)) - \tfrac{1}{n}\mathbf{1}_n\big)^2 \\
= \sum_{i=1}^n \sum_{j=1}^n \big(\hat{v}_{ij}(t_1^\ell(+)) - \hat{v}_{jj}(t_0^{\ell_1 j})\big)^2 , \\
\leq n^2 \big(\xi/n^2\big) ,
\end{aligned}
$$

where the first equality holds again due to Lemma 7.9. ∎

Proof.(Theorem 4.7) Lemma 7.20-7.22. *Overview of Proof.* The Lemma 7.20 proves the OH normal consensus update (22). As stated in Lemma 7.21, the update (22) implies the error of each element of every normal consensus vector is a non-increasing function of time, and thus (22) will imply the conditions stated in Lemma 7.22 are sufficient and necessary for any node to obtain average-consensus. The Theorem 4.7 then follows immediately from Lemma 7.22.

Lemma 7.20: (OH Normal Consensus Estimate Update) Applying (20) and (21) to the optimization problem (9) yields the OH normal consensus estimate update $\hat{v}_i(t_1^{ij}(+))$ defined in (22).

Proof: If $\hat{v}_j(t_0^{ij}) = \frac{1}{n}\mathbf{1}_n$ then $0 = S_1^{ij}(t_0^{ij}, t_1^{ij})$, in this case (20) together with (8) implies (48). The update (9) can thus be re-written as (49). Since $\hat{s}_j(t_0^{ij}) = \bar{s}(0)$ is known we can let $\hat{s}_i(t_1^{ij}(+)) = \bar{s}(0)$ and thus obtain $\hat{v}_i(t_1^{ij}(+)) = \frac{1}{n}\mathbf{1}_n$ as the unique solution to (49). Note that this coincides with (22). Next assume that $\hat{v}_j(t_0^{ij}) \neq \frac{1}{n}\mathbf{1}_n$ and thus $0 \neq S_1^{ij}(t_0^{ij}, t_1^{ij})$. Under (20) and (21) we can re-write (9) as,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) = \quad &\arg_{\tilde{v}} \ \min \ \big(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\big)^2, \\
&\text{s.t. (8) holds, given } \hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij}), \\
&s_j(0) = \mathbf{S}e_j, \ s_i(0) = \mathbf{S}e_i.
\end{aligned}
\tag{113}
$$

Given that $\hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij})$, $s_j(0) = \mathbf{S}e_j$, and $s_i(0) = \mathbf{S}e_i$ are known, the set of vectors $\hat{v}_i(t_1^{ij}(+))$ for which (8) holds is $\text{span}\{\hat{v}_i(t_1^{ij}), e_j, e_i\}$, thus (113) can be re-written as,

$$
\hat{v}_i(t_1^{ij}(+)) = \arg_{\tilde{v}} \min_{\tilde{v} \in \text{span}\{\hat{v}_i(t_1^{ij}), e_j, e_i\}} \big(\tilde{v} - \frac{1}{n}\mathbf{1}_n\big)^2.
\tag{114}
$$

If $\hat{v}_i(t_1^{ij}) \in \text{span}\{e_j, e_i\}$ then (114) becomes,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) \quad &= \arg_{\tilde{v}} \ \min_{\tilde{v} \in \text{span}\{e_j, e_i\}} \big(\tilde{v} - \tfrac{1}{n}\mathbf{1}_n\big)^2 , \\
&= V_{(OH1)} V_{(OH1)}^+ \tfrac{1}{n}\mathbf{1}_n , \quad V_{(OH1)} = [e_j, e_i] .
\end{aligned}
$$

Since e_i is linearly independent of e_j we then have,

$$
\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) &= V_{(OH1)}(V_{(OH1)}' V_{(OH1)})^{-1} V_{(OH1)}' \tfrac{1}{n}\mathbf{1}_n \\
&= \tfrac{1}{n} V_{(OH1)}(I_2)^{-1}\mathbf{1}_2 , \\
&= \tfrac{1}{n}(e_i + e_j) .
\end{aligned}
\tag{115}
$$

Note that if $\hat{v}_i(t_1^{ij}) \in \text{span}\{e_j, e_i\}$ then the last line in (115) coincides with (22). Next assume $\hat{v}_i(t_1^{ij}) \notin \text{span}\{e_j, e_i\}$. In this case (114) can be expressed,

$$\hat{v}_i(t_1^{ij}(+)) = V_{(OH)}V_{(OH)}^+ \frac{1}{n}\mathbf{1}_n \ ,$$
$$= V_{(OH)}(V_{(OH)}'V_{(OH)})^{-1}V_{(OH)}'\frac{1}{n}\mathbf{1}_n$$
$$V_{(OH)} = [\hat{v}_i(t_1^{ij}), \tfrac{1}{n}e_j, \tfrac{1}{n}e_i] \ .$$

Recall the discrete set of vectors $\mathbb{R}_{0,\frac{1}{n}}^n$ is defined in (25). Note that the initialization (45) implies $\hat{v}_{ii}(0) = \frac{1}{n}$ and thus $\hat{v}_i(0) \in \mathbb{R}_{0,\frac{1}{n}}^n$. Also note that (115) implies $\hat{v}_{ii}(t_1^{ij}(+)) = \frac{1}{n}$ and $\hat{v}_i(t_1^{ij}(+)) \in \mathbb{R}_{0,\frac{1}{n}}^n$. We thus assume,

$$\hat{v}_{ii}(t_1^{ij}) = \frac{1}{n} \ , \quad \hat{v}_i(t_1^{ij}) \in \mathbb{R}_{0,\frac{1}{n}}^n \ . \tag{116}$$

Observe that under (116) if the result (22) is proven, then the assumptions (116) are valid. For notational convenience denote $\hat{v}_i = \hat{v}_i(t_1^{ij})$. Given (116) the matrix $(V_{(OH)}'V_{(OH)})$ has an inverse (117) below,

$$
\begin{aligned}
(V_{(OH)}'V_{(OH)})^{-1} \ &= n^{-4}\begin{pmatrix} \hat{v}_i'\mathbf{1}_n & 1 & v_{ij} \\ 1 & 1 & 0 \\ v_{ij} & 0 & 1 \end{pmatrix}^{-1} \\
&= \frac{n^{-4}}{|V_{(OH)}'V_{(OH)}|}\begin{pmatrix} 1 & -1 & -v_{ij} \\ -1 & (\hat{v}_i'\mathbf{1}_n - v_{ij}) & v_{ij} \\ -v_{ij} & v_{ij} & (\hat{v}_i'\mathbf{1}_n - 1) \end{pmatrix}
\end{aligned}
\tag{117}
$$

where the determinant $|V_{(OH)}'V_{(OH)}|$ can be computed as,

$$|V_{(OH)}'V_{(OH)}| = n^{-6}(\hat{v}_i'\mathbf{1}_n - 1 - v_{ij}) \ .$$

Next observe that,

$$V_{(OH)}'\frac{1}{n}\mathbf{1}_n = \frac{1}{n^2}\begin{pmatrix} \hat{v}_i'\mathbf{1}_n \\ 1 \\ 1 \end{pmatrix} \ . \tag{118}$$

From (118) we observe that right-multiplying (117) by $V_{(OH)}'\frac{1}{n}\mathbf{1}_n$ and left-multiplying by $V_{(OH)}$ then yields (22),

$$
\begin{aligned}
V_{(OH)}&(V_{(OH)}'V_{(OH)})^{-1}V_{(OH)}'\frac{1}{n}\mathbf{1}_n \\
&= \frac{1}{\hat{v}_i'\mathbf{1}_n - 1 - \hat{v}_{ij}}V_{(OH)}\begin{pmatrix} \hat{v}_i'\mathbf{1}_n - 1 - \hat{v}_{ij} \\ 0 \\ \hat{v}_i'\mathbf{1}_n(1 - \hat{v}_{ij}) + \hat{v}_{ij} - 1 \end{pmatrix} \\
&= \frac{1}{n}v^{ij}(t_1^{ij}, t_0^{ij}) \ .
\end{aligned}
$$

$\blacksquare$

Lemma 7.21: (OH Element-Wise Non-Increasing Error of the Normal Consensus Estimate) The error $\left(\hat{v}_{i\ell}(t) - 1/n\right)^2$ of each element $\hat{v}_{i\ell}(t)$ is a non-increasing function of $t \geq 0$ for all $(i, \ell) \in \mathcal{V}^2$.

Proof: The result follows immediately from the normal consensus update (22). $\blacksquare$

Lemma 7.22: (OH Normal Consensus Estimate Convergence to Average-Consensus) Under the OH algorithm, any node $i \in \mathcal{V}$ obtains average-consensus by time t_1 for some communication sequence $C_{[0,t_1)}$ iff at least one of the following two conditions holds,

- (C1i): there is a signal $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[0,t_1)}$ for each $j \in \mathcal{V}_{-i}$.
- (C2i): there is a communication path $C_{[t_0(i\ell), t_1(i\ell)]}^{i\ell} \in C_{[0,t_1)}$ from at least one node ℓ such that $S^{\ell j}(t_0^{\ell j}, t_1^{\ell j}) \in C_{[0,t_0(i\ell))}$ for all $j \in \mathcal{V}_{-\ell}$.

Proof: (Sufficiency.) If a communication sequence $C_{[0,t_1)}$ implies (C1i) for some node i then under the update (22) there will exist a time $t_1^{ij} \in [0, t_1)$ such that $\hat{v}_{ij}(t_1^{ij}(+)) = \frac{1}{n}$ for each $j \in \mathcal{V}_{-i}$, and thus Lemma 7.21 together with (45) imply $\hat{v}_i(t_1) = \frac{1}{n}\mathbf{1}_n$ and hence by (8) the node i reaches average-consensus by time t_1. If a communication sequence $C_{[0,t_1)}$ implies (C2i) then by the previous reasoning the node ℓ will obtain average-consensus by time $t_0(i\ell)$, and thus by the update (22) any node $j \in \mathcal{V}$ that ℓ sends a signal $S^{j\ell}(t_0^{j\ell}, t_1^{j\ell}) \in C_{[t_0(i\ell), t_1)}$ will obtain average-consensus at $t_1^{j\ell}(+)$. If the node ℓ has a communication path $C_{[t_0(i\ell), t_1(i\ell)]}^{i\ell} \in C_{[0,t_1)}$ to node i it then follows that node i will have obtained average-consensus by time $t_1(i\ell)(+)$, and thus by Lemma 7.21 node i obtains average-consensus at time t_1.

(Necessity.) Under the OH update (22), if there is not a signal $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[0,t_1)}$ then there will not exist a time $t \in [0, t_1)$ such that $\hat{v}_{ij}(t) = \frac{1}{n}$ unless there exists a communication path $C_{[t_0(i\ell), t_1(i\ell)]}^{i\ell} \in C_{[0,t_1)}$ from some node ℓ such that $S^{\ell j}(t_0^{\ell j}, t_1^{\ell j}) \in C_{[0,t_0(i\ell))}$ for all $j \in \mathcal{V}_{-\ell}$. It thus follows that node i cannot obtain average-consensus by time t_1 for any communication sequence $C_{[0,t_1)}$ that does not imply either (C1i) or (C2i). ∎

Theorem 4.7 (OH Network Convergence to Average-Consensus)

Proof: The condition (C) stated in Theorem 4.7 is equivalent to when either (C1i) or (C2i) hold for each node $i \in \mathcal{V}$, the result thus follows immediately from Lemma 7.22. ∎

Proof. (Theorem 4.9) Lemma 7.23-7.35. *Overview of Proof.* Similar to Theorem 4.5, we show that every normal consensus estimate $\hat{v}_i(t)$ satisfies the normalization property (54), the "zero local error" property (64), and furthermore the discretization $\hat{v}_i(t) \in \mathbb{R}_{0,\frac{1}{n}}^n$. The Lemma 7.28 and Lemma 7.29 shows that the update (123) respectively implies the error of each normal consensus estimate is non-decreasing with time, and that the normal consensus estimate will not change unless there is a reduction in error. We then show, analogous to Lemma 7.13, that the reduction in error that results from any signal will eventually vanish if $C_{[0,t_1]}$ satisfies (37). Due to the discretization $\hat{v}_i(t) \in \mathbb{R}_{0,\frac{1}{n}}^n$ this implies the reduction in error that results from any signal will eventually strictly equal zero if $C_{[0,t_1]}$ satisfies (37), see Lemma 7.35. When this occurs we can show that $\hat{v}_i(t) = \frac{1}{n}\mathbf{1}_n$ by utilizing Lemma 7.30, Lemma 7.31, together with the "zero local error" property (64), hence (44) holds and so by Definition 7.1 a network average-consensus is obtained.

Lemma 7.23: (DDA Normal Consensus Estimate Discretization) Every normal consensus estimate $\hat{v}_i(t)$ satisfies $\hat{v}_i(t) \in \mathbb{R}_{0,\frac{1}{n}}^n$ for all $i \in \mathcal{V}$ and $t \geq 0$.

Proof: Note that the initialization (45) implies $\hat{v}_i(0) \in \mathbb{R}_{0,\frac{1}{n}}^n$. The optimization problem (26) requires that any solution $\hat{v}_i(t(+))$ satisfies $\hat{v}_i(t(+)) \in \mathbb{R}_{0,\frac{1}{n}}^n$. Under the DDA algorithm, the assumption (A7) implies every normal

consensus estimate remains fixed unless updated via (26), it thus follows that $\hat{v}_i(t) \in \mathbb{R}^n_{0,\frac{1}{n}}$ for all $i \in \mathcal{V}$ and $t \geq 0$. ∎

Lemma 7.24: (DDA Consensus Estimate Normalization) Every normal consensus estimate $\hat{v}_i(t)$ satisfies (54), and furthermore,

$$\left(\hat{v}_i^{-i}(t)\right)^2 = \frac{1}{n}\hat{v}_i^{-i}(t)'\mathbf{1}_{n-1} \ , \ \forall \ i \in \mathcal{V} \ , \ \forall \ t \geq 0 \ . \tag{119}$$

Proof: The Lemma 7.23 implies,

$$\begin{aligned}
\tfrac{1}{n}\hat{v}_i(t)'\mathbf{1}_n \ &= \tfrac{1}{n}\sum_{\ell=1}^n \hat{v}_{i\ell}(t) \ , \\
&= \tfrac{1}{n}\sum_{\ell \in \mathcal{V} \ : \ \hat{v}_{i\ell}(t)=\frac{1}{n}} \left(\tfrac{1}{n}\right) \ , \\
&= \sum_{\ell \in \mathcal{V} \ : \ \hat{v}_{i\ell}(t)=\frac{1}{n}} \left(\tfrac{1}{n^2}\right) \ , \\
&= \sum_{\ell=1}^n \hat{v}_{i\ell}(t)^2 \ , \\
&= \hat{v}_i(t)^2 \ .
\end{aligned} \tag{120}$$

If $\hat{v}_i(t) \in \mathbb{R}^n_{0,\frac{1}{n}}$ then $\hat{v}_i^{-i}(t) \in \mathbb{R}^{n-1}_{0,\frac{1}{n}}$, thus a similar argument to (120) implies (119). ∎

Lemma 7.25: (DDA Local Zero Error Property) Every normal consensus estimate $\hat{v}_i(t)$ satisfies (64).

Proof: The Lemma 7.2 implies $\hat{v}_{ii}(0) = \frac{1}{n}$ for each $i \in \mathcal{V}$. Next observe that under $(A7)$ the estimate $\hat{v}_i(t)$ will not change unless a signal $S^{ij}(t_0^{ij}, t_1^{ij})$ is received at node i. If a signal is received then $\hat{v}_i(t)$ is updated by (123) given below. We now show that under the assumption (64), the solution $\hat{v}_i(t_1^{ij}(+))$ specified by (123) will imply (64) for every set vectors $\{e_i, \hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij})\}$. If $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) = 0$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i$ and thus,

$$\begin{aligned}
\hat{v}_{ii}(t_1^{ij}(+)) \ &= \hat{v}_{ii}(t_1^{ij}) + \hat{v}_{ji}(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij}) \ , \\
&= \tfrac{1}{n} \ .
\end{aligned}$$

If $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$ and $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_j(t_0^{ij}) + e_i\left(\frac{1}{n} - \hat{v}_{ji}(t_0^{ij})\right)$ and thus,

$$\begin{aligned}
\hat{v}_i(t_1^{ij}(+)) \ &= \hat{v}_{ji}(t_0^{ij}) + \tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij}) \ , \\
&= \tfrac{1}{n} \ .
\end{aligned}$$

Finally, if $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$ and $\hat{v}_i^{-i}(t_1^{ij})^2 \geq \hat{v}_j^{-i}(t_0^{ij})^2$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij})$ and thus (64) follows by assumption. ∎

Lemma 7.26: (DDA Normal Consensus Estimate Magnitude Equivalence to Error) For any two normal consensus estimates $\hat{v}_i(t)$ and $\hat{v}_j(t)$,

$$\hat{v}_i(t)^2 \geq \hat{v}_j(t)^2 \ \Leftrightarrow \ E^2_{\frac{1}{n}\mathbf{1}_n}\left(\hat{v}_i(t)\right) \leq E^2_{\frac{1}{n}\mathbf{1}_n}\left(\hat{v}_j(t)\right) \ . \tag{121}$$

Proof: For any normal consensus estimate $\hat{v}_i(t)$,

$$\begin{aligned}
E^2_{\frac{1}{n}\mathbf{1}_n}\left(\hat{v}_i(t)\right) \ &= \left(\hat{v}_i(t) - \tfrac{1}{n}\mathbf{1}_n\right)^2 \ , \\
&= \hat{v}_i(t)^2 - 2\tfrac{1}{n}\hat{v}_i(t)'\mathbf{1}_n + \tfrac{1}{n} \ , \\
&= \tfrac{1}{n} - \hat{v}_i(t)^2
\end{aligned} \tag{122}$$

where the third line holds due to Lemma 7.24. The equivalence (121) then follows directly from (122). ∎

Lemma 7.27: (DDA Normal Consensus Estimate Update) Applying (27) and (28) to the optimization problem (26) implies the DDA normal consensus estimate update $\hat{v}_i(t_1^{ij}(+))$ can be defined as in $(29) - (30)$.

Proof: Note that $(29) - (30)$ implies,

$$\hat{v}_i(t_1^{ij}(+)) = \begin{cases} \hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) = 0, \\ \hat{v}_j(t_0^{ij}) + e_i\left(\frac{1}{n} - \hat{v}_{ji}(t_0^{ij})\right) & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0 \, , \ \hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2 \, , \\ \hat{v}_i(t_1^{ij}) & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0 \, , \ \hat{v}_i^{-i}(t_1^{ij})^2 \geq \hat{v}_j^{-i}(t_0^{ij})^2 \, . \end{cases} \tag{123}$$

We thus will prove that (27) and (28) imply optimization problem (26) yields the DDA normal consensus estimate update $\hat{v}_i(t_1^{ij}(+))$ defined by (123). Under (27) and (28) we can re-write (26) as,

$$\begin{aligned} \hat{v}_i(t_1^{ij}(+)) = \ & \arg_{\tilde{v}} \ \min \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n\right)^2, \\ & \text{s.t. (8) holds}, \tilde{v} \in \mathbb{R}^n_{0,\frac{1}{n}}, \text{ given } \hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij}), \\ & \hat{s}_j(t_0^{ij}) = \mathbf{S}\hat{v}_j(t_0^{ij}), s_i(0) = \mathbf{S}e_i. \end{aligned} \tag{124}$$

Given that $\hat{s}_i(t_1^{ij}) = \mathbf{S}\hat{v}_i(t_1^{ij}), \hat{s}_j(t_0^{ij}) = \mathbf{S}\hat{v}_j(t_0^{ij})$, and $s_i(0) = \mathbf{S}e_i$ are known, the set of vectors $\hat{v}_i(t_1^{ij}(+))$ for which (8) holds is $\text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\}$, thus (124) can be re-written as

$$\begin{aligned} \hat{v}_i(t_1^{ij}(+)) \ &= \arg_{\tilde{v}} \ \min_{\tilde{v} \, \in \, \text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\} \cap \mathbb{R}^n_{0,\frac{1}{n}}} \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n\right)^2, \\ &= \arg_{\tilde{v} \, = \, \left(a\hat{v}_i(t_1^{ij}) + b\hat{v}_j(t_0^{ij}) + c\frac{1}{n}e_i\right) \cap \mathbb{R}^n_{0,\frac{1}{n}}} \ \min_{(a,b,c)} \ \left(\tilde{v} - \frac{1}{n}\mathbf{1}_n\right)^2 . \end{aligned} \tag{125}$$

For notational convenience denote $\hat{v}_i = \hat{v}_i(t_1^{ij})$, $\hat{v}_j = \hat{v}_j(t_0^{ij})$, and $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(+)$. Note that the constraint in (125) can be expressed as follows,

$$a\hat{v}_{i\ell} + b\hat{v}_{j\ell} + c\frac{1}{n}e_{i\ell} \in \left(0, \frac{1}{n}\right) \, , \ \forall \, \ell \in \mathcal{V}. \tag{126}$$

Due to Lemma 7.25 the ith constraint in (126) is,

$$\begin{aligned} & a\frac{1}{n} + b\hat{v}_{ji} + c\frac{1}{n} \in \left(0, \frac{1}{n}\right) \, , \ \Rightarrow \\ & c = \begin{cases} -a + b\hat{v}_{ji}n \, , \\ 1 - a - b\hat{v}_{ji}n \, . \end{cases} \end{aligned}$$

If $c = -a + b\hat{v}_{ji}n$ we then have the candidate update $\hat{v}_i^{(1)}(+)$,

$$\begin{aligned} \hat{v}_i^{(1)}(+) \ &= a\hat{v}_i + b\hat{v}_j + \frac{1}{n}e_i(-a + b\hat{v}_{ji}n) \, , \ \Rightarrow \\ \hat{v}_i^{(1)}(+)^2 \ &= \frac{1}{n}\hat{v}_i^{(1)}(+)'\mathbf{1}_n \\ &= a\hat{v}_i^2 + b\hat{v}_j^2 - a\frac{1}{n^2} - b\hat{v}_{ji}/n \, , \end{aligned}$$

where the magnitude $\hat{v}_i^{(1)}(+)^2$ is computed using Lemma 7.24. If $c = 1 - a + b\hat{v}_{ji}n$ we then have the candidate update $\hat{v}_i^{(2)}(+)$,

$$\begin{aligned} \hat{v}_i^{(2)}(+) \ &= a\hat{v}_i + b\hat{v}_j + \frac{1}{n}e_i(1 - a + b\hat{v}_{ji}n) \, , \ \Rightarrow \\ \hat{v}_i^{(2)}(+)^2 \ &= \frac{1}{n}\hat{v}_i^{(1)}(+)'\mathbf{1}_n \\ &= a\hat{v}_i^2 + b\hat{v}_j^2 + \frac{1}{n^2} - a\frac{1}{n^2} - b\hat{v}_{ji}/n \, , \\ &= \hat{v}_i^{(1)}(+)^2 + \frac{1}{n^2} \, . \end{aligned} \tag{127}$$

Since $e_{i\ell} = 0$ for all $\ell \neq i$, the ith constraint of (126 is the only constraint involving the optimization variable c, thus by applying Lemma 7.26 it follows from (127) that if $c \neq 1 - a + b\hat{v}_{ji}n$ then the resulting solution $\hat{v}_i(+)$ cannot be a solution to (125), hence $\hat{c} = 1 - a + b\hat{v}_{ji}n$. Next observe that if $\hat{v}_{i\ell} = \hat{v}_{j\ell} = 0$ then the ℓth constraint in (126) places no restriction on a or b, thus due to Lemma 7.23 and Lemma 7.25 we can consider the three possible scenarios posed by (126) given the ith constraint is satisfied by (127): one constraint,

$$(M1A) \quad a(0) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (0,1) \;,$$
$$(M1B) \quad a(\tfrac{1}{n}) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (1-a, -a) \;,$$

two constraints,

$$(M2A) \quad a(0) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (0,1)$$
$$\qquad\qquad a(\tfrac{1}{n}) + b(0) \in (0, \tfrac{1}{n}) \;\Rightarrow\; a \in (0,1) \;,$$
$$(M2B) \quad a(0) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (0,1)$$
$$\qquad\qquad a(\tfrac{1}{n}) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; a \in (1-b, -b) \;,$$
$$(M2C) \quad a(\tfrac{1}{n}) + b(0) \in (0, \tfrac{1}{n}) \;\Rightarrow\; a \in (0,1)$$
$$\qquad\qquad a(\tfrac{1}{n}) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (1-a, -a) \;,$$

or three constraints

$$(M3) \quad a(0) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (0,1)$$
$$\qquad\quad a(\tfrac{1}{n}) + b(0) \in (0, \tfrac{1}{n}) \;\Rightarrow\; a \in (0,1)$$
$$\qquad\quad a(\tfrac{1}{n}) + b(\tfrac{1}{n}) \in (0, \tfrac{1}{n}) \;\Rightarrow\; b \in (1-a, -a) \;.$$

If $\hat{v}_i^{-i}{}'\hat{v}_j^{-i} = 0$ and there is one constraint, then Lemma 7.25 implies $\hat{v}_i^{-i} = 0$ and hence $(M1A)$. In this case we have $\hat{v}_i = \tfrac{1}{n}e_i$ and the following candidate solutions,

$$\hat{v}_i(+) = \begin{cases} a\tfrac{1}{n}e_i + \tfrac{1}{n}e_i(1-a) = \tfrac{1}{n}e_i, & (b=0) \\ a\tfrac{1}{n}e_i + \hat{v}_j + \tfrac{1}{n}e_i(1 - a - \hat{v}_{ji}n), & (b=1) \end{cases} \tag{128}$$

where the second line in (128) simplifies to $\hat{v}_j + \tfrac{1}{n}e_i(1 - \hat{v}_{ji}n)$. The solutions in (128) possess the following magnitudes,

$$\hat{v}_i(+)^2 = \begin{cases} \tfrac{1}{n^2} \;, & (b=0) \\ \hat{v}_j^{-i}{}^2 + \tfrac{1}{n^2} \;, & (b=1) \end{cases}$$

and thus since Lemma 7.25 implies $\hat{v}_j^{-i}{}^2 > 0$, the optimal solution is,

$$\begin{aligned} \hat{v}_i(+) &= \hat{v}_j + \tfrac{1}{n}e_i(1 - \hat{v}_{ji}n) \;, \\ &= \hat{v}_i + \hat{v}_j - e_i\hat{v}_{ji} \;, \end{aligned} \tag{129}$$

where the last line follows under the given assumption $\hat{v}_i = \tfrac{1}{n}e_i$. If $\hat{v}_i^{-i}{}'\hat{v}_j^{-i} = 0$ and there are two constraints then $(M2A)$ necessarily follows. In this case $\hat{v}_i^{-i}{}^2 > 0$ and we have the candidate solutions,

$$\hat{v}_i(+) = \begin{cases} \tfrac{1}{n}e_i \;, & (a=0,\ b=0) \\ \hat{v}_i \;, & (a=1,\ b=0) \\ \hat{v}_j + \tfrac{1}{n}e_i(1 - \hat{v}_{ji}n) \;, & (a=0,\ b=1) \\ \hat{v}_i + \hat{v}_j - e_i\hat{v}_{ji} \;, & (a=1,\ b=1) \;. \end{cases} \tag{130}$$

The solutions in (130) possess the following magnitudes,

$$
\hat{v}_i(+)^2 = \begin{cases}
\frac{1}{n^2} \ , & (a=0,\ b=0) \\
\hat{v}_i^{-i\ 2} + \frac{1}{n^2} \ , & (a=1,\ b=0) \\
\hat{v}_j^{-i\ 2} + \frac{1}{n^2} \ , & (a=0,\ b=1) \\
\hat{v}_i^{-i\ 2} + \hat{v}_j^{-i\ 2} + \frac{1}{n^2} \ , & (a=1,\ b=1)
\end{cases}
$$

and thus since Lemma 7.25 implies $\hat{v}_j^{-i\ 2} > 0$, the optimal solution is,

$$
\hat{v}_i(+) = \hat{v}_i + \hat{v}_j - e_i \hat{v}_{ji} \ . \tag{131}
$$

If $\hat{v}_i^{-i}{}' \hat{v}_j^{-i} > 0$ and there is one constraint then we necessarily have $\hat{v}_i^{-i} = \hat{v}_j^{-i}$ and thus $(M1B)$. In this case the candidate solutions are,

$$
\hat{v}_i(+) = \begin{cases}
a\hat{v}_i + (1-a)\hat{v}_j + \frac{1}{n}e_i(1 - a - (1-a)\hat{v}_{ji}n) \\
a\hat{v}_i - a\hat{v}_j + \frac{1}{n}e_i(1 - a + a\hat{v}_{ji}n) \ .
\end{cases} \tag{132}
$$

Note that the first and second line in (132) correspond respectively to when $b = 1 - a$ and $b = -a$, also note that each line simplifies respectively to $\hat{v}_j + \frac{1}{n}e_i(1 - \hat{v}_{ji}n)$ and $\frac{1}{n}e_i$. The solutions in (132) thus possess the following magnitudes,

$$
\hat{v}_i(+)^2 = \begin{cases}
\hat{v}_j^{-i\ 2} + \frac{1}{n^2} \ , & (b = 1 - a) \\
\frac{1}{n^2} \ , & (b = -a) \ ,
\end{cases}
$$

and since Lemma 7.25 implies $\hat{v}_j^{-i\ 2} > 0$, the optimal solution is,

$$
\begin{aligned}
\hat{v}_i(+) &= \hat{v}_j + \frac{1}{n}e_i(1 - \hat{v}_{ji}n) \ , \\
&= \hat{v}_i \ ,
\end{aligned} \tag{133}
$$

where the last line follows under the assumption $\hat{v}_i^{-i} = \hat{v}_j^{-i}$. If $\hat{v}_i^{-i}{}' \hat{v}_j^{-i} > 0$ and there are two constraints, then if $\hat{v}_i^{-i\ 2} < \hat{v}_j^{-i\ 2}$ it follows that $(M2B)$ holds. The candidate solutions in this case are,

$$
\hat{v}_i(+) = \begin{cases}
\frac{1}{n}e_i \ , & (a=0,\ b=0) \\
\hat{v}_i \ , & (a=1,\ b=0) \\
-\hat{v}_i + \hat{v}_j - \frac{1}{n}e_i(2 - \hat{v}_{ji}n) \ , & (a=-1,\ b=1) \\
\hat{v}_j + \frac{1}{n}e_i(1 - \hat{v}_{ji}n) \ , & (a=0,\ b=1).
\end{cases} \tag{134}
$$

The solutions in (134) possess the following magnitudes,

$$
\hat{v}_i(+)^2 = \begin{cases}
\frac{1}{n^2} \ , & (a=0,\ b=0) \\
\hat{v}_i^{-i\ 2} + \frac{1}{n^2} \ , & (a=1,\ b=0) \\
-\hat{v}_i^{-i\ 2} + \hat{v}_j^{-i\ 2} + \frac{1}{n^2} \ , & (a=-1,\ b=1) \\
\hat{v}_j^{-i\ 2} + \frac{1}{n^2} \ , & (a=0,\ b=1).
\end{cases}
$$

The assumption $\hat{v}_j^{-i\ 2} > \hat{v}_i^{-i\ 2}$ then implies the optimal solution,

$$
\hat{v}_i(+) = \hat{v}_j - \frac{1}{n}e_i(1 - \hat{v}_{ji}n) \ . \tag{135}
$$

If $\hat{v}_i^{-i}\,{}'\hat{v}_j^{-i} > 0$ and there are two constraints, then if $\hat{v}_i^{-i} \geq \hat{v}_j^{-i}$ it follows that $(M2C)$ holds. The candidate solutions in this case are,

$$\hat{v}_i(+) = \begin{cases} \frac{1}{n}e_i\,, & (a=0,\ b=0) \\ \hat{v}_j + \frac{1}{n}e_i(1-\hat{v}_{ji}n)\,, & (a=0,\ b=1) \\ \hat{v}_i - \hat{v}_j + \frac{1}{n}e_i\,, & (a=1,\ b=-1) \\ \hat{v}_i\,, & (a=1,\ b=0)\,. \end{cases} \tag{136}$$

The solutions in (136) possess the following magnitudes,

$$\hat{v}_i(+)^2 = \begin{cases} \frac{1}{n^2}\,, & (a=0,\ b=0) \\ \hat{v}_j^{-i\,2} + \frac{1}{n^2}\,, & (a=0,\ b=1) \\ \hat{v}_i^{-i\,2} - \hat{v}_j^{-i\,2} + \frac{1}{n^2}\,, & (a=1,\ b=-1) \\ \hat{v}_i^{-i\,2} + \frac{1}{n^2}\,, & (a=1,\ b=0). \end{cases}$$

The assumption $\hat{v}_j^{-i\,2} \leq \hat{v}_i^{-i\,2}$ then implies the global solution,

$$\hat{v}_i(+) = \hat{v}_i\,. \tag{137}$$

If $\hat{v}_i^{-i}\,{}'\hat{v}_j^{-i} > 0$ and there are three constraints then $(M3)$ necessarily follows. The candidate solutions in this case are,

$$\hat{v}_i(+) = \begin{cases} \frac{1}{n}e_i\,, & (a=0,\ b=0) \\ \hat{v}_j + \frac{1}{n}e_i(1-\hat{v}_{ji}n)\,, & (a=0,\ b=1) \\ \hat{v}_i\,, & (a=1,\ b=0). \end{cases} \tag{138}$$

The solutions in (138) possess the following magnitudes,

$$\hat{v}_i(+)^2 = \begin{cases} \frac{1}{n^2}\,, & (a=0,\ b=0) \\ \hat{v}_j^{-i\,2} + \frac{1}{n^2}\,, & (a=0,\ b=1) \\ \hat{v}_i^{-i\,2} + \frac{1}{n^2}\,, & (a=1,\ b=0). \end{cases}$$

The assumption $\hat{v}_j^{-i\,2} > \hat{v}_i^{-i\,2}$ then implies the optimal solution,

$$\hat{v}_i(+) = \hat{v}_j + \frac{1}{n}e_i(1-\hat{v}_{ji}n)\,. \tag{139}$$

In contrast, the assumption $\hat{v}_j^{-i\,2} \leq \hat{v}_i^{-i\,2}$ implies the global solution,

$$\hat{v}_i(+) = \hat{v}_i\,. \tag{140}$$

Combining (129), (131), (133), (135), (137), (139), and (140), together imply (123). Note that if $\hat{v}_j^{-i\,2} = \hat{v}_i^{-i\,2}$ and $\hat{v}_j^{-i} \neq \hat{v}_i^{-i}$ then there are two global solutions to (125). By specifying (137) and (140) we have chosen the global solutions that are necessary for Lemma 7.29, which is in turn necessary for the DDA algorithm to obtain average-consensus under the sufficient communication condition stated in Theorem 4.9. ∎

Lemma 7.28: (DDA Normal Consensus Estimate Non-Decreasing Error) The error of every normal consensus estimate $\hat{v}_i(t)$ is a non-decreasing function of $t \geq 0$ for all $i \in \mathcal{V}$.

Proof: The Lemma 7.27 implies that upon reception of any signal $S^{ij}(t_0^{ij}, t_1^{ij})$, the normal consensus estimate $\hat{v}_i(t_1^{ij})$ is updated to a solution of (125), thus

$$\hat{v}_i(t_1^{ij}(+)) \in \text{span}\{\hat{v}_i(t_1^{ij}), \hat{v}_j(t_0^{ij}), e_i\} \cap \mathbb{R}_{0,\frac{1}{n}}^n . \tag{141}$$

The Lemma 7.23 implies $\hat{v}_i(t_1^{ij}) \in \mathbb{R}_{0,\frac{1}{n}}^n$, thus (141) implies that any candidate solution $\hat{v}^{(1)}(t_1^{ij}(+))$ that does not satisfy $E_{\frac{1}{n}\mathbf{1}_n}(\hat{v}^{(1)}(t_1^{ij}(+))) \leq E_{\frac{1}{n}\mathbf{1}_n}(\hat{v}(t_1^{ij}))$ cannot be a solution (125). ∎

Lemma 7.29: (DDA Fixed Normal Consensus Estimate Given No Reduction in Error) Under the DDA algorithm, for any signal $S^{ij}(t_0^{ij}, t_1^{ij})$ we have,

$$E_{\frac{1}{n}}^2(\hat{v}_i(t_1^{ij}(+))) = E_{\frac{1}{n}}^2(\hat{v}_i(t_1^{ij})) \Rightarrow \hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij}). \tag{142}$$

Proof: Applying Lemma 7.26 to the LHS of (142) implies,

$$E_{\frac{1}{n}}^2(\hat{v}_i(t_1^{ij}(+))) = E_{\frac{1}{n}}^2(\hat{v}_i(t_1^{ij}))$$
$$\Leftrightarrow (v_i(t_1^{ij}(+)))^2 = (\hat{v}_i(t_1^{ij}))^2 .$$

To prove (142) we thus have only to show,

$$\hat{v}_i(t_1^{ij}(+)) \neq \hat{v}_i(t_1^{ij}) \Rightarrow (\hat{v}_i(t_1^{ij}(+)))^2 \neq (\hat{v}_i(t_1^{ij}))^2 . \tag{143}$$

Under the update (123), to prove (143) we need only to show that either of the two cases,

$$\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i ,$$
$$\text{if } \hat{v}_i^{-i}(t_1^{ij})' \hat{v}_j^{-i}(t_0^{ij}) = 0,$$
$$\hat{v}_i(t_1^{ij}(+)) = \hat{v}_j(t_0^{ij}) + e_i(\tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij})) , \tag{144}$$
$$\text{if } \hat{v}_i^{-i}(t_1^{ij})' \hat{v}_j^{-i}(t_0^{ij}) > 0, \hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$$

will imply $\hat{v}_i^2(t_1^{ij}(+)) > \hat{v}_i^2(t_1^{ij})$. If $\hat{v}_i^{-i}(t_1^{ij})' \hat{v}_j^{-i}(t_0^{ij}) = 0$ then (144) together with (54) imply,

$$\begin{aligned}
\hat{v}_i^2(t_1^{ij}(+)) &= \tfrac{1}{n}\hat{v}_i(t_1^{ij}(+))'\mathbf{1}_n , \\
&= \tfrac{1}{n}(\hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i)'\mathbf{1}_n , \\
&= \tfrac{1}{n}(\hat{v}_i(t_1^{ij})'\mathbf{1}_n + \hat{v}_j(t_0^{ij})'\mathbf{1}_n - \hat{v}_{ji}(t_0^{ij})) , \\
&= \hat{v}_i^2(t_1^{ij}) + \tfrac{1}{n}\hat{v}_j^{-i}(t_0^{ij})'\mathbf{1}_{n-1} , \\
&> \hat{v}_i^2(t_1^{ij}) ,
\end{aligned}$$

where the last inequality holds because (64) implies $\hat{v}_j^{-i}(t)'\mathbf{1}_n \geq \tfrac{1}{n}$ for all $j \in \mathcal{V}_{-i}$ and $t \geq 0$. If $\hat{v}_i^{-i}(t_1^{ij})' \hat{v}_j^{-i}(t_0^{ij}) > 0$ and $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$, then due to (144) it follows that,

$$\begin{aligned}
\hat{v}_i^2(t_1^{ij}(+)) &= \tfrac{1}{n}\hat{v}_i(t_1^{ij}(+))'\mathbf{1}_n , \\
&= \tfrac{1}{n}(\hat{v}_j(t_0^{ij}) + e_i(\tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij})))'\mathbf{1}_n , \\
&= \tfrac{1}{n}(\hat{v}_j(t_0^{ij})'\mathbf{1}_n + \tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij})) , \\
&= \tfrac{1}{n}(\hat{v}_j^{-i}(t_0^{ij})'\mathbf{1}_{n-1} + \tfrac{1}{n}) , \\
&> \tfrac{1}{n}(\hat{v}_i^{-i}(t_1^{ij})'\mathbf{1}_{n-1} + \tfrac{1}{n}) , \\
&= \tfrac{1}{n}\hat{v}_i(t_1^{ij})'\mathbf{1}_n = \hat{v}_i^2(t_1^{ij}) ,
\end{aligned}$$

where the final inequality holds under the given assumption that $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$, and the second last equality holds due to Lemma 7.25. ∎

Lemma 7.30: (DDA Lower Bound on Increase in Normal Consensus Magnitude) Under the DDA algorithm, for any signal $S^{ij}(t_0^{ij}, t_1^{ij})$ we have,

$$\left(\hat{v}_i^{-i}(t_1^{ij}(+))\right)^2 \geq \max\{\left(\hat{v}_i^{-i}(t_1^{ij})\right)^2 , \ \left(\hat{v}_j^{-i}(t_0^{ij})\right)^2\} .$$

Proof: If $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) = 0$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i$ and thus applying (119) implies,

$$\begin{aligned}
\left(\hat{v}_i^{-i}(t_1^{ij}(+))\right)^2 &= \tfrac{1}{n}\hat{v}_i^{-i}(t_1^{ij}(+))'\mathbf{1}_{n-1} , \\
&= \tfrac{1}{n}\left(\hat{v}_i^{-i}(t_1^{ij})'\mathbf{1}_{n-1} + \hat{v}_j^{-i}(t_0^{ij})'\mathbf{1}_{n-1}\right) , \\
&\geq \max\{\left(\hat{v}_i^{-i}(t_1^{ij})\right)^2 , \ \left(\hat{v}_j^{-i}(t_0^{ij})\right)^2\} .
\end{aligned}$$

If $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$ and $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_j(t_0^{ij}) + e_i\left(\tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij})\right)$ and thus,

$$\begin{aligned}
\left(\hat{v}_i^{-i}(t_1^{ij}(+))\right)^2 &= \tfrac{1}{n}\left(\hat{v}_j^{-i}(t_0^{ij}) + e_i^{-i}\left(\tfrac{1}{n} - \hat{v}_{ji}(t_0^{ij})\right)\right)\mathbf{1}_{n-1} \\
&= \tfrac{1}{n}\hat{v}_j^{-i}(t_0^{ij})'\mathbf{1}_{n-1} \\
&\geq \max\{\left(\hat{v}_i^{-i}(t_1^{ij})\right)^2 , \ \left(\hat{v}_j^{-i}(t_0^{ij})\right)^2\}
\end{aligned}$$

where the last inequality holds by the assumption $\hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2$. Finally, if $\hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0$ and $\hat{v}_i^{-i}(t_1^{ij})^2 \geq \hat{v}_j^{-i}(t_0^{ij})^2$ then $\hat{v}_i(t_1^{ij}(+)) = \hat{v}_i(t_1^{ij})$ and thus

$$\begin{aligned}
\left(\hat{v}_i^{-i}(t_1^{ij}(+))\right)^2 &= \tfrac{1}{n}\hat{v}_i^{-i}(t_1^{ij}(+))'\mathbf{1}_{n-1} , \\
&= \tfrac{1}{n}\hat{v}_i^{-i}(t_1^{ij})'\mathbf{1}_{n-1} , \\
&\geq \max\{\left(\hat{v}_i^{-i}(t_1^{ij})\right)^2 , \ \left(\hat{v}_j^{-i}(t_0^{ij})\right)^2\} ,
\end{aligned}$$

where the last inequality holds by the assumption $\hat{v}_i^{-i}(t_1^{ij})^2 \geq \hat{v}_j^{-i}(t_0^{ij})^2$. ∎

Lemma 7.31: (DDA Upper Bound on the Magnitude $\left(\hat{v}_i^{-i}(t)\right)^2$) If $\hat{v}_i^2(t) = \hat{v}_i^2(t')$ for all $(t, t') \in [t_0, t_1]$ then

$$\left(\hat{v}_i^{-i}(t)\right)^2 \leq \left(\hat{v}_i^{-j}(t')\right)^2 , \ \forall \ j \in \mathcal{V}_{-i} , \ \forall \ (t, t') \in [t_0, t_1].$$

Proof: The result follows since $\hat{v}_i(t) \in \mathbb{R}^n_{0, \frac{1}{n}}$ implies $\hat{v}_{ij}(t) \in (0, \tfrac{1}{n})$ for all $j \in \mathcal{V}$ and $t \geq 0$, thus,

$$\begin{aligned}
\left(\hat{v}_i^{-i}(t)\right)^2 &= \hat{v}_i^2(t) - \tfrac{1}{n^2} , \\
&= \hat{v}_i^2(t') - \tfrac{1}{n^2} , \\
&\leq \hat{v}_i^2(t') - \hat{v}_{ij}^2(t') , \\
&= \left(\hat{v}_i^{-j}(t')\right)^2 .
\end{aligned}$$

∎

Lemma 7.32: (Error Expression for $C_{[0, t_1]}$ satisfying (37)) For any communication sequence $C_{[0, t_1]}$ satisfying (37) the total reduction in normal consensus error from $t = 0$ to $t = t_1(+)$ is,

$$\begin{aligned}
\mathbf{E}^2(C_{[0, t_1]}) &= \sum_{i=1}^n \left(E_{\frac{1}{n}\mathbf{1}_n}^2\left(\hat{v}_i(0)\right) - E_{\frac{1}{n}\mathbf{1}_n}^2\left(\hat{v}_i(t_1(+))\right)\right) \\
&= \tfrac{n-1}{n} - \sum_{i=1}^n E_{\frac{1}{n}\mathbf{1}_n}^2\left(\hat{v}_i(t_1(+))\right) \\
&= \sum_{\ell \in \mathbb{N}} \mathbf{E}^2(C_{[t_0^\ell, t_1^\ell]}) , \ C_{[t_0^\ell, t_1^\ell]} \in \mathcal{SVCC} \\
&\leq \tfrac{n-1}{n} .
\end{aligned} \tag{145}$$

Proof: The proof is identical to Lemma 7.12 with (86), (35), and S$\mathcal{V}$SC replaced by (145), (37), and S$\mathcal{V}$CC respectively. ∎

Lemma 7.33: (DDA Vanishing Reduction in Error for $C_{[0,t_1]}$ satisfying (37)) For any communication sequence $C_{[0,t_1]}$ satisfying (37) there exists an integer ℓ_ε such that,

$$\mathbf{E}^2(C_{[t_0^\ell,t_1^\ell]}) \le \varepsilon \ , \ \forall \ \ell \ge \ell_\varepsilon \ ,$$

for any $\varepsilon > 0$.

Proof: The proof is identical to Lemma 7.13 with (86) and (35) replaced by (145) and (37) respectively. ∎

Lemma 7.34: (DDA Lower Bound on Non-Zero Reduction in Error) For any communication sequence $C_{[t_0,t_1]}$ we have,

$$\mathbf{E}^2(C_{[t_0,t_1]}) > 0 \ \Rightarrow \ \mathbf{E}^2(C_{[t_0,t_1]}) \ge \frac{1}{n^2} \ . \tag{146}$$

Proof: Applying Lemma 7.26 to the LHS of (146) implies there exists some signal $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[t_0,t_1]}$ such that,

$$\hat{v}_i^2(t_1^{ij}(+)) > \hat{v}_i^2(t_1^{ij}) \ . \tag{147}$$

We now show that (147) implies $\hat{v}_i^2(t_1^{ij}(+)) \ge \hat{v}_i^2(t_1^{ij}) + \frac{1}{n^2}$, the result (146) then follows directly by Lemma 7.28 together with Lemma 7.26. The Lemma 7.23 implies $\hat{v}_i(t) \in \mathbb{R}^n_{0,\frac{1}{n}}$ for all $i \in \mathcal{V}$ and $t \ge 0$, thus under (147) it follows that,

$$\begin{aligned}
\hat{v}_i^2(t_1^{ij}(+)) &= \tfrac{1}{n}\hat{v}_i(t_1^{ij}(+))'\mathbf{1}_n \ , \\
&= \tfrac{1}{n^2}|\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij}(+))\}| \ , \\
&> \hat{v}_i^2(t_1^{ij}) \ , \\
&= \tfrac{1}{n}\hat{v}_i(t_1^{ij})'\mathbf{1}_n \ , \\
&= \tfrac{1}{n^2}|\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij})\}| \ .
\end{aligned} \tag{148}$$

From (148) it then follows,

$$\begin{aligned}
|\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij}(+))\}| &> |\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij})\}| \ \Rightarrow \\
|\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij}(+))\}| - 1 &\ge |\{\ell \ : \ \hat{v}_{i\ell}(t_1^{ij})\}| \ .
\end{aligned} \tag{149}$$

Under the constraint $\hat{v}_i(t) \in \mathbb{R}^n_{0,\frac{1}{n}}$ the last inequality in (149) implies $\hat{v}_i^2(t_1^{ij}(+)) \ge \hat{v}_i^2(t_1^{ij}) + \frac{1}{n^2}$. ∎

Lemma 7.35: (DDA Existence of a Time for Zero Reduction in Error) For any communication sequence $C_{[0,t_1]}$ satisfying (37) there exists an integer $\ell_{\frac{1}{n^2}}$ such that,

$$\mathbf{E}^2(C_{[t_0^\ell,t_1^\ell]}) = 0 \ , \ \forall \ \ell \ge \ell_{\frac{1}{n^2}} \ . \tag{150}$$

Proof: The result follows immediately by applying Lemma 7.28 and Lemma 7.34 to Lemma 7.33. ∎

Theorem 4.9(DDA Network Convergence to Average-Consensus)

Proof: For any communication sequence $C_{[0,t_1]}$ satisfying (37) the Lemma 7.35 implies there exists an integer $\ell_{\frac{1}{n^2}}$ such that (150) holds. We now show that the condition,

$$\mathbf{E}^2(C_{[t_0^\ell,t_1^\ell]}) = 0 \ , \ C_{[t_0^\ell,t_1^\ell]} \in S\mathcal{V}CC \tag{151}$$

implies (44) at $t = t_0^\ell$, and hence by Definition 7.1 a network average-consensus is obtained at t_0^ℓ. Applying Lemma 7.30 to (151) implies,

$$
\begin{aligned}
\left(\hat{v}_{\hat{i}}^{-\hat{i}}(t_1^{\hat{i}j}(+))\right)^2 &\geq \left(\hat{v}_j^{-\hat{i}}(t_0^{\hat{i}j})\right)^2 , \ \forall \, j \in \mathcal{V}_{-\hat{i}} , \\
&\geq \left(\hat{v}_j^{-j}(t_0^{ij})\right)^2 , \ \forall \, j \in \mathcal{V}_{-\hat{i}}
\end{aligned}
\tag{152}
$$

where the last inequality holds due to Lemma 7.31. Likewise, applying Lemma 7.30 and Lemma 7.31 to (151) implies,

$$
\begin{aligned}
\left(\hat{v}_j^{-j}(t_1^{j\hat{i}}(+))\right)^2 &\geq \left(\hat{v}_{\hat{i}}^{-j}(t_0^{j\hat{i}})\right)^2 , \ \forall \, j \in \mathcal{V}_{-\hat{i}} , \\
&\geq \left(\hat{v}_{\hat{i}}^{-\hat{i}}(t_0^{j\hat{i}})\right)^2 , \ \forall \, j \in \mathcal{V}_{-\hat{i}} .
\end{aligned}
\tag{153}
$$

Note that applying Lemma 7.29 to (151) implies that (152) and (153) can be combined to obtain,

$$
\begin{aligned}
\left(\hat{v}_{\hat{i}}^{-\hat{i}}(t_0^\ell)\right)^2 &= \left(\hat{v}_{\hat{i}}^{-j}(t_0^\ell)\right)^2 , \ \forall \, j \in \mathcal{V}_{-\hat{i}} \ \Rightarrow \\
\hat{v}_{\hat{i}}^2(t_0^\ell) - \tfrac{1}{n^2} &= \hat{v}_{\hat{i}}^2(t_0^\ell) - \hat{v}_{\hat{i}j}^2(t_0^\ell) , \ \forall \, j \in \mathcal{V}_{-\hat{i}} .
\end{aligned}
\tag{154}
$$

The second line in (154) together with Lemma 7.25 implies,

$$
\hat{v}_{\hat{i}j}(t_0^\ell) = \frac{1}{n} , \ \forall \, j \in \mathcal{V} ,
$$

and thus $\hat{v}_{\hat{i}}(t_0^\ell) = \frac{1}{n}\mathbf{1}_n$. Since $C_{[t_0^\ell, t_1^\ell]} \in \text{S}\mathcal{V}\text{CC}$, the Lemma 7.29 together with the update (123) implies $\hat{v}_j(t_0^\ell) = \frac{1}{n}\mathbf{1}_n$ for each $j \in \mathcal{V}_{-\hat{i}}$, thus (44) holds and so by Definition 7.1 average-consensus is obtained at time $t_0^\ell < t_1$. ∎

Remark 7.36: We observe that if the S$\mathcal{V}$CC condition is defined by the condition (C) stated in Theorem 4.7, then using the definition (37) for the I$\mathcal{V}$CC condition will not imply Theorem 4.9. In other words, using condition (C) to define an S$\mathcal{V}$CC sequence will imply there exist examples of I$\mathcal{V}$CC sequences for which the DDA algorithm will not obtain average-consensus. This is why we have defined S$\mathcal{V}$CC only by (36), which is actually a special case of the condition (C) stated in Theorem 4.7. Furthermore, the DDA algorithm normal consensus update (123) is only a global solution to (26), it is not a unique solution. Under (20) and (21), the alternative global solution to (26) is,

$$
\hat{v}_i(t_1^{ij}(+)) = \begin{cases}
\hat{v}_i(t_1^{ij}) + \hat{v}_j(t_0^{ij}) - \hat{v}_{ji}(t_0^{ij})e_i & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) = 0 , \\
\hat{v}_j(t_0^{ij}) + e_i\left(\frac{1}{n} - \hat{v}_{ji}(t_0^{ij})\right) & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0 , \ \hat{v}_i^{-i}(t_1^{ij})^2 < \hat{v}_j^{-i}(t_0^{ij})^2 , \\
\hat{v}_i(t_1^{ij}) & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0 , \ \hat{v}_i^{-i}(t_1^{ij})^2 > \hat{v}_j^{-i}(t_0^{ij})^2 , \\
\hat{v}_j(t_0^{ij}) & , \text{ if } \hat{v}_i^{-i}(t_1^{ij})'\hat{v}_j^{-i}(t_0^{ij}) > 0 , \ \hat{v}_i^{-i}(t_1^{ij})^2 = \hat{v}_j^{-i}(t_0^{ij})^2 .
\end{cases}
\tag{155}
$$

The above remark still holds even when the alternative global solution (155) is used to update the normal consensus estimate; however, by randomly switching between the updates (123) and (155) leads to the following conjecture.

Conjecture 7.37: Let the S$\mathcal{V}$CC condition be defined by the condition (C) stated in Theorem 4.7. Suppose upon reception of each signal the normal consensus estimate update is defined by (123) with probability $p \in (0, 1)$ and defined by (155) with probability $1 - p$. Then (6) holds at time $t = t_1(+)$ almost surely for any communication sequence $C_{[0, t_1]}$ satisfying (37).

The significance of the above result is due to the fact that condition (C) in Theorem 4.7 is considerably weaker than (36). If Conjecture 7.37 holds, then by defining the DDA algorithm using the above randomized protocol, and defining the S$\mathcal{V}$CC condition by the condition (C) stated in Theorem 4.7, the Venn diagram in Fig.1 will then be

completely accurate. We note there are alternative definitions of the S$\mathcal{V}$CC condition and OH algorithm that will also lead to the same Venn diagram as Fig.1. However, we know of no *weaker* sufficient condition than that stated in Theorem 4.9 for convergence of the DDA algorithm $(27) - (32)$, and this sufficient condition is based entirely on the S$\mathcal{V}$CC condition (36).

B. Comparison Algorithms: Gossip and ARIS

In this section we define the two comparison algorithms, Gossip and ARIS, in terms of the class of distributed algorithms $(3), (7)$.

1) Comparison Algorithm 1 (Gossip): The Gossip algorithm proposed in [3] implies a signal specification and knowledge set update defined respectively as (156) and (157) below.

Gossip Algorithm:

Signal Specification:
$$S^{ij}(t_0^{ij}, t_1^{ij}) = \mathcal{K}_j(t_0^{ij}) \tag{156}$$

Knowledge Set Update:
$$\mathcal{K}_i(t_1^{ij}(+)) = \{\hat{s}_i(t_1^{ij}(+))\}$$
$$\hat{s}_i(t_1^{ij}(+)) = \tfrac{1}{2}\big(\hat{s}_i(t_1^{ij}) + \hat{s}_j(t_0^{ij})\big) \tag{157}$$

Initialization:
$$\mathcal{K}_i(0) = \{\hat{s}_i(0)\} \ , \ \ \hat{s}_i(0) \ = s_i(0) \ , \ \forall \ i \in \mathcal{V} \ . \tag{158}$$

We note that under the Gossip algorithm the only communication conditions proven to ensure average-consensus require instantaneous and bi-directional communication, thus implying

$$S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[0,t_1]} \Leftrightarrow S^{ji}(t_0^{ji}, t_1^{ji}) \in C_{[0,t_1]} \ ,$$
$$t_0^{ij} = t_1^{ij} = t_0^{ji} = t_1^{ji} \ , \tag{159}$$

for any signal $S^{ij}(t_0^{ij}, t_1^{ij}) \in C_{[0,t_1]}$. Under the assumption (159), any $C_{[0,t_1]}$ satisfying the I$\mathcal{V}$SC condition (35) will imply the Gossip algorithm obtains average-consensus at time $t = t_1(+)$, that is (6) is satisfied at $t = t_1(+)$. We note that in some works (e.g. [17], [2]) the Gossip algorithm is referred to as "pairwise averaging".

2) Comparison Algorithm 2 (ARIS): The adaptation of the randomized information spreading algorithm proposed in [19] that we will consider can be defined by the signal specification (160) and knowledge set update (161) below.

ARIS Algorithm:

Signal Specification:
$$S^{ij}(t_0^{ij}, t_1^{ij}) = \mathcal{K}_j(t_0^{ij}) \setminus \{s_i(0), i, n\} \tag{160}$$

Knowledge Set Update:
$$\mathcal{K}_i(t_1^{ij}(+)) = \{k_i(t_1^{ij}(+)), \hat{s}_i(t_1^{ij}(+)),$$
$$W^i(t_1^{ij}(+)), \hat{w}_i(t_1^{ij}(+)), s_i(0), i, n\} \tag{161}$$

$$\text{Initialization:} \qquad \begin{aligned} &\mathcal{K}_i(0) = \{k_i(0), \hat{s}_i(0), W^i(0), \hat{w}_i(0), s_i(0), i, n\} \\ &k_i(0) = 0 \ , \ \hat{s}_i(0) = s_i(0)/n \ , \ \hat{w}_i(0) = e_i \ , \\ &W^i(0) \in \mathbb{R}^{d \times r} \ , \ W^i_{\ell q}(0) \ \sim \ \exp\{s_{i\ell}(0)\} \ , \\ &\forall \ \ell = 1, 2, \ldots, d \ , \ q = 1, 2, \ldots, r \ . \end{aligned} \qquad (162)$$

We clarify that the $(\ell q)^{th}$ element in the matrix $W^i(0)$ is an independent realization of a random variable from an exponential distribution with rate parameter $s_{i\ell}(0)$, this is why the elements of each initial consensus vector $s_i(0)$ are required to be positive valued under any version of the RIS algorithm. We next define the ARIS update for each term in the knowledge set $\mathcal{K}_i(t_1^{ij}(+))$, we will omit the time indices for convenience.

ARIS Knowledge Set Update Procedure:

$$\hat{w}^1 = \begin{cases} \mathbf{1}_n - \delta[\hat{w}_i + \hat{w}_j], & \text{if } k_j = k_i, \\ \mathbf{1}_n - \delta[e_i + \hat{w}_j], & \text{if } k_j > k_i, \\ 0, & \text{if } k_j < k_i, \end{cases}$$

$$k_i(+) = \begin{cases} k_j, & \text{if } k_j > k_i, \hat{w}^1 \neq \mathbf{1}_n \\ k_j + 1, & \text{if } k_j > k_i, \hat{w}^1 = \mathbf{1}_n \\ k_i, & \text{if } k_j = k_i, \hat{w}^1 \neq \mathbf{1}_n \\ k_i + 1, & \text{if } k_j = k_i, \hat{w}^1 = \mathbf{1}_n \\ k_i, & \text{if } k_j < k_i, \end{cases} \qquad (163)$$

$$\hat{w}_i(+) = \begin{cases} \hat{w}^1, & \text{if } \hat{w}^1 \notin \{\mathbf{1}_n, 0\} \\ e_i, & \text{if } \hat{w}^1 = \mathbf{1}_n \\ \hat{w}_i, & \text{if } \hat{w}^1 = 0, \end{cases} \qquad (164)$$

$$\hat{W}^1_{\ell q} = \begin{cases} \sim \exp\{s_{i\ell}(0)\}, & \text{if } k_i(+) > k_i \\ \min\{W^j_{\ell q}, W^i_{\ell q}\}, & \text{if } k_i(+) = k_i = k_j \\ W^i_{\ell q}, & \text{if } k_j < k_i \end{cases}$$
$$\ell = 1, 2, \ldots, d, \quad q = 1, 2, \ldots, r,$$

$$W^i_{\ell q}(+) = \begin{cases} \sim \exp\{s_{i\ell}(0)\}, & \text{if } k_j > k_i, \hat{w}^1 = \mathbf{1}_n \\ \min\{W^j_{\ell q}, \hat{W}^1_{\ell q}\}, & \text{if } k_j = k_i, \hat{w}^1 \neq \mathbf{1}_n \\ \hat{W}^1_{\ell q}, & \text{if } k_j \leq k_i \end{cases} \qquad (165)$$
$$\ell = 1, 2, \ldots, d, \quad q = 1, 2, \ldots, r,$$

$$\bar{w}_{i\ell} = \frac{1}{r} \sum_{q=1}^r \min\{W^j_{\ell q}(t_0^{ij}), W^i_{\ell q}(t_1^{ij})\} \ , \ \ell = 1, 2, \ldots, d,$$

$$\underline{w}_{i\ell} = \frac{1}{r} \sum_{q=1}^r \min\{W^j_{\ell q}(t_0^{ij}), \hat{W}^1_{\ell q}\} \quad , \ \ell = 1, 2, \ldots, d.$$

$$\hat{s}_i(+) = \begin{cases} \hat{s}_i, & \text{if } k_j < k_i, \\[2mm] \hat{s}_i, & \text{if } k_j = k_i, \hat{w}^1 \neq \mathbf{1}_n \\[2mm] \frac{k_i \hat{s}_i + \bar{w}_i^{-1}/n}{k_i+1}, & \text{if } k_j = k_i, \hat{w}^1 = \mathbf{1}_n \\[2mm] \hat{s}_j, & \text{if } k_j > k_i, \hat{w}^1 \neq \mathbf{1}_n \\[2mm] \frac{k_j \hat{s}_j + \underline{\mathbf{W}}_i^{-1}/n}{k_j+1}, & \text{if } k_j > k_i, \hat{w}^1 = \mathbf{1}_n \end{cases} \tag{166}$$

The integer r is an RIS algorithm parameter that affects the convergence error of the algorithm; as r increases the algorithm is expected to converge closer to the true average-consensus vector $\bar{s}(0)$. Based on the strong law of large numbers we make the following conjecture.

Conjecture 7.38: For any $C_{[0,t_1]}$ satisfying the S$\mathcal{V}$SC condition (34), both the RIS algorithm proposed in [19] and the ARIS algorithm $(160) - (166)$ imply (6) at time $t = t_1(+)$ almost surely in the limit as r approaches infinity.

As explained in Sec.V-A, the total resource cost of the RIS algorithm increases on the order $O(rd)$, and likewise, the total resource cost of the ARIS algorithm increases on the order $O(rd + n)$, thus it is not practical to assume r can be made arbitrarily large as the Conjecture 7.38 requires. Note that due to the resource costs of the RIS and ARIS algorithms, the Conjecture 7.38 does not contradict Conjecture 7.4, even though both assume the same communication conditions.

An informal description of the ARIS algorithm is as follows: at $t = 0$ each node generates r random variables independently from an exponential distribution with rate parameter $s_{i\ell}(0)$ for each $\ell = 1, 2, \ldots, d$. A local counter k_i is set to zero and a local vector $\hat{w}_i$ is set to e_i. Upon reception of any signal, if the transmitted counter k_j is less than the local counter k_i, then the signal is ignored, if $k_j = k_i$ then the receiving node records the minimum between the received r random variables and local r random variables for each $\ell = 1, 2, \ldots, d$. The vector $\hat{w}_i$ maintains a record of the set of nodes that have a communication path with the node i for any given counter value. Whenever $\hat{w}_i(t)$ is updated to $\mathbf{1}_n$, then the counter k_i is increased by one, the consensus estimate $\hat{s}_i(t)$ is updated as a running average of the inverse of the currently recorded mean of the minimum random variables for each $\ell = 1, 2, \ldots, d$, $\hat{w}_i(t)$ is reset to e_i, and a new set of rd random variables are locally generated. If $k_j > k_i$ then the local counter k_i is set to k_j, the consensus estimate $\hat{s}_i(t)$ is set to the received estimate $\hat{s}_j$, $\hat{w}_i(t)$ is reset to e_i and updated by $\hat{w}_j$, a new set of rd random variables are locally generated, and the node i then records the minimum between the received r random variables and the newly generated local r random variables for each $\ell = 1, 2, \ldots, d$. It is not difficult to see that the local counter of each node will approach infinity iff $C_{[0,t_1]}$ satisfies the I$\mathcal{V}$SC condition (35), thus each element in the consensus estimate $\hat{s}_i(t)$ of each node $i \in \mathcal{V}$ will approach the inverse of the mean of the minimum of infinitely many random variables. By a well-known property of the minimum of a set of independently generated exponential random variables, together with the strong law of large numbers, we thus make the following conjecture.

Conjecture 7.39: For any positive integer r, the ARIS algorithm $(160) - (166)$ implies (6) at time $t = t_1(+)$ iff the communication sequence $C_{[0,t_1]}$ satisfies the IVSC condition (35).

The above conjecture is significant because, besides the flooding technique and the DA algorithm $(15) - (19)$, there is no other consensus protocol in the literature that guarantees average-consensus for every communication sequence $C_{[0,t_1]}$ that satisfies the IVSC condition (35).

C. Resource Cost Derivations

In this section we explain how the entries of Table I in Sec.V-A are obtained. We will assume an arbitrary vector $v \in \mathbb{R}^m$ requires $2m$ scalars to be defined, and similarly, an unordered set of scalars S with cardinality $|S|$ requires $|S|$ scalars to be defined, that is

$$\text{Resource cost of } v \in \mathbb{R}^m = 2m,$$
$$\text{Resource cost of } S = |S|. \tag{167}$$

The rationale for (167) is that each element in v requires one scalar to define the location of the element, and one scalar to define the value of the element itself. The location of each scalar in S is irrelevant because S is an unordered set, thus S can be defined using only $|S|$ scalars. An alternative to (167) is to assume that an arbitrary vector $v \in \mathbb{R}^m$ requires m scalars to be defined. Although this alternative will imply different entries for the *exact* values in Table I, the *order* of the storage and communication costs under each algorithm would remain the same. We adhere to (167) for our resource cost computations due to its relative precision.

Next observe that under the BM, OH, and DDA algorithms each normal consensus estimate $\hat{v}_i(t)$ will contain only elements belonging to the set $\{0, \frac{1}{n}\}$, similarly each vector $\hat{w}_i(t)$ under the ARIS algorithm contains only elements belonging to the set $\{0, 1\}$. Under the BM, OH, and DDA algorithms we can thus define each $\hat{v}_i(t)$ based only on the set $\hat{V}_i(t)$,

$$\hat{V}_i(t) = \begin{cases} \{-\ell \ : \ \hat{v}_{i\ell}(t) = 0\} & \text{if } n\hat{v}_i(t)'\mathbf{1}_n \geq 1/2, \\ \{ \ \ell \ : \ \hat{v}_{i\ell}(t) = \frac{1}{n}\} & \text{if } n\hat{v}_i(t)'\mathbf{1}_n < 1/2. \end{cases} \tag{168}$$

A set $\hat{W}_i(t)$ can be defined analogous to (168) to specify $\hat{w}_i(t)$. Note that both $\hat{V}_i(t)$ and $\hat{W}_i(t)$ are unordered sets of scalars, and thus, assuming average-consensus has not been obtained, we have from (168), (45), and (162),

$$1 \leq \{|\hat{V}_i(t)|, |\hat{W}_i(t)|\} \leq \lfloor n/2 \rfloor . \tag{169}$$

Under the condition (45), applying (167) and (169) to the algorithm update equations (10), (11), (20), (21), (27), (28), then yield the respective upper and lower resource costs for the BM, OH, and DDA algorithms stated in Table I in Sec.V-A. Likewise, applying (167) and (169) to (160)-(165) yield the respective upper and lower resource costs for the ARIS algorithm. As detailed in $(156) - (158)$, the Gossip algorithm proposed in [3] requires only the local consensus estimate $\hat{s}_i(t)$ to be communicated and stored at each node, thus the communication and storage cost are both fixed at $2d$ under this algorithm.

We next observe that under the DA algorithm $(15) - (19)$, if a normal consensus estimate $\hat{v}_i(t)$ contains any elements equal to zero then these elements may be omitted from the signal and knowledge set. Given the condition

(45) it follows from (167) that the minimum number of scalars required to define $\hat{v}_i(t)$ under the DA algorithm is 2, while the maximum number of scalars required to define $\hat{v}_i(t)$ is $2n$. Together with these upper and lower limits on the resource cost of $\hat{v}_i(t)$, applying (167) to (15) and (16) then yields the upper and lower resource costs for the DA algorithm stated in Table I in Sec.V-A.

D. Example SVSC Sequence

The sequence defined in (170) below is an SVSC sequence that implies the DA, DDA, and BM algorithm all obtain average-consensus at the same instant.

$$
\begin{aligned}
&C_{[0,4n-5]} = \{S_1, S_2, \ldots, S_{2(n-1)}\} \ , \\
&S_i = S^{i+1,i}(2(i-1), 2i-1) \ , \ i = 1, 2, \ldots, n-1, \\
&S_n = S^{1,n}(2(n-1), 2n-1) \ , \\
&S_i = S^{i-n+1,i-n}(2(i-1), 2i-1) \ , \ i = n+1, n+2, \ldots, 2n-2 \ .
\end{aligned}
\tag{170}
$$

It is a "unit-delay" sequence because each signal sent at time t is received at time $t+1$, and if a signal is received at time $t+1$ then the next signal is sent at time $t+2$. Together with Theorem 4.3, the example (170) implies that the DA and DDA algorithms possess the weakest possible necessary conditions for average-consensus that any algorithm can have. In contrast, the OH algorithm does not achieve average-consensus under (170).

REFERENCES

[1] M. Bawa, H. Garcia-Molina, A. Gionis, and R. Motwani, *Estimating Aggregates on a Peer-to-Peer Network*, Technical Report, Stanford University, 2003. URL: http://dbpubs.stanford.edu/pub/2003-24.

[2] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsiklis, *Convergence in Multi-agent Coordination, Consensus, and Flocking*, in Proceedings of IEEE Conf. on Decision and Control, Seville, Spain, 2005.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, *Randomized Gossip Algorithms*, IEEE Trans. Info. Theory, vol.52, pp. 2508-2530, 2006.

[4] S. Boyd, L. Xiao, and S. Lall, *A Scheme for Robust Distributed Sensor Fusion based on Average Consensus*, in Proceedings of the 4th International IEEE Symposium on Information Processing in Sensor Networks, Los Angeles, CA, 2005.

[5] S. Boyd, L. Xiao, and S. Lall, *Distributed Average Consensus with Time-varying Metropolis Weights*, in Proceedings of the 4th International Conference on Information Processing in Sensor Networks, Los Angeles, CA, 2005.

[6] J. Considine, F. Li, G. Kollios, and J. Byers, *Approximate Aggregation Techniques for Sensor Databases*, International Conference on Data Engineering, 2004.

[7] J. Cortes, *Finite-time Convergent Gradient Flows with Applications to Network Consensus*, Automatica, vol.42, no.11, pp. 1993-2000, 2006.

[8] G. Cybenko, *Dynamic Load Balancing for Distributed Memory Multiprocessors*, Journal of Parallel and Distributed Computing, vol.7, no.2, pp.279-301, 1989.

[9] M. Franceschelli, A. Giua, and C. Seatzu, *Consensus on the Average on Arbitrary Strongly Connected Digraphs Based on Broadcast Gossip Algorithms*, 1st IFAC Workshop on Estimation and Control of Networked Systems, Venice, Italy, 2009.

[10] Y. Hatano and M. Mesbahi, *Agreement over Random Networks*, IEEE Trans. on Automat. Control, vol.50, no.11, pp.1867-1872, 2005.

[11] S. Kar and J. Moura, *Distributed Consensus Algorithms in Sensor Networks with Imperfect Communication: Link Failures and Channel Noise*, IEEE Transactions on Signal Processing, vol.57, no.1, pp.355-369, 2009.

[12] D. Kempe, A. Dobra, and J. Gehrke, *Gossip-based Computation of Aggregate Information*, in Proceedings of the 44th IEEE Symposium on Foundations of Computer Science, pp. 482-491, 2003.

[13] V. Krishnamurthy, K. Topley, and G. Yin, *Consensus Formation in a Two-Time-Scale Markovian System*, SIAM MMS, vol.7, no.4, pp.1898-1927, 2009.

[14] T. Li and J. Zhang, *Mean Square Average-Consensus under Measurement Noises and Fixed Topologies: Necessary and Sufficient Conditions*, Automatica, vol.45, pp. 1929-1936, 2009.

[15] P. Lin and Y. Jia, *Average Consensus in Networks of Multi-agent with both Switching Topology and Coupling Time-Delay*, Physica A, vol.387, pp.303-313, 2008.

[16] M. Mehyar, D. Spanos, J. Pongsajapan, S. Low, and R. Murray, *Asynchronous Distributed Averaging on Communication Networks*, IEEE Trans. on Networking, vol.15, no.3, pp.512-520, 2007.

[17] C. Moallemi and B. Van Roy, *Consensus Propagation*, IEEE Trans. on Info. Theory, vol.54, no.7, pp.2997-3007, 2008.

[18] L. Moreau, *Stability of Multi-agent Systems with Time-dependent Communication Links*, IEEE Trans. on Automat. Control, vol.50, no.2, pp.169-182, 2005.

[19] D. Mosk-Aoyama and D. Shah, *Fast Gossip Algorithm for Computing Separable Functions*, IEEE Trans. on Info. Theory, vol.54, no.7, pp.2997-3007, 2008.

[20] R. Olfati-Saber, *Distributed Kalman Filter with Embedded Consensus Filters*, in Proceedings of the 44th IEEE Conf. on Decision and Control, Seville, Spain, 2005.

[21] R. Olfati-Saber and R. Murray, *Consensus Problems in Networks of Agents with Switching Topology and Time-Delays*, IEEE Trans. on Automatic Control, vol.49, no.9, 2004.

[22] R. Olfati-Saber, R. Murray, and D. Spanos, *Distributed Sensor Fusion using Dynamic Consensus*, in Proceedings of the 16th IFAC World Congress, Prague, Czech Republic, 2005.

[23] M. Porfiri and D. Stilwell, *Consensus Seeking Over Random Weighted Directed Graphs*, IEEE Trans. on Automatic Control, vol.52, no.9, pp. 1767-1773, 2007.

[24] W. Ren and R. Beard, *Consensus Seeking in Multi-Agent Systems under Dynamically Changing Interaction Topologies*, IEEE Trans. on Automatic Control, vol.50, no.5, pp.655-661, 2005.

[25] A. Tahbaz-Salehi and A. Jadbabaie, *A Necessary and Sufficient Condition for Consensus Over Random Networks*, IEEE Trans. on Automatic Control, vol.53, no.3, 2008.

[26] B. Touri and A. Nedic, *Distributed Consensus over Network with Noisy Links*, 12th International Conference on Information Fusion, Seattle, WA, 2009.

[27] J. Tsitsiklis, D. Bertsekas, and M. Athans, *Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms*, IEEE Trans. on Automatic Control, vol.31, no.9, pp.803-812, 1986.

[28] T. Zhang and H. Yu, *Average Consensus in Networks of Multi-agent with Multiple Time-varying Delays*, in International Journal of Communications and System Sciences, vol.3, pp.196-203, 2010.

[29] M. Zhu and S. Martinez, *Dynamic Average Consensus on Synchronous Communication Networks*, in Proceedings of the American Control Conference, Seattle, 2008.